THE IMPACT OF POVERTY ON COMPARABLE IMPROVEMENT RANKING FOR ELEMENTARY CAMPUSES IN TEXAS

Vicki Gay Holland, B.S., M.Ed.

Dissertation Prepared for the Degree of

DOCTOR OF EDUCATION

UNIVERSITY OF NORTH TEXAS

December 2001

APPROVED:

Frank R. Kemerer, Major Professor
Patricia A. Moseley, Minor Professor
Robin K. Henson, Committee Member
Millie Gore, Committee Member
Jane B. Huffman, Program Coordinator
M. Jean Keller, Dean of College of Education
C. Neal Tate, Dean of the Robert B. Toulouse School of Graduate Studies

Holland, Vicki Gay, The impact of comparable improvement ranking for elementary campuses in Texas. Doctor of Education (Educational Administration), December 2001, 141 pp., 3 tables, 7 figures, references, 140 titles.

The problem was to determine how comparable is comparable improvement for campuses in Texas. An alternative strategy for determining comparable improvement was developed using 2000 comparable improvement data provided by the Texas Education Agency for 2,403 elementary campuses. Comparable improvement is a measure that shows how student performance has changed from one year to the next and then compares that growth to 40 schools that are demographically most similar to the target school. Instead of using the most dominant characteristic as in the current process, the percent of students in poverty was the initial sorting characteristic. The impact of sorting by poverty was reviewed in four areas: 1.) the impact on quartile placement, 2.) the TLI average growth for the comparison group, 3.) the award eligibility, and 4.) the changes in comparison group composition.

No practical significant difference was found for research questions 1 and 3, however, a practical significant difference was found in group average TLI growth for math and in the comparison group composition. Overall, the alternative process had the greatest impact on campuses with 40-80% poverty. Three possible factors may have influenced the results. First, the middle poverty campuses had the most change in comparison group as found in question 4. Second, the interaction between the middle poverty campuses and the alternative process could have been fueled by the removal of

the 1,295 campuses with poverty as the dominant characteristic in current system. Third, the high correlation between poverty and ethnicity may have limited any impact of the alternative process.

ACKNOWLEDGEMENTS

Sincere appreciation and admiration is extended to Dr. Frank Kemerer, my major professor, for his continuous guidance and leadership in pursuing and completing my degree. Without his knowledge and support, this project could not have been successfully completed.

A special note of appreciation is given to Dr. Robin Henson for his assistance in statistical design and analysis. Also, appreciation is extended to the supportive graduate committee members, Dr. Patricia Moseley and Dr. Millie Gore, for their assistance in completing this dissertation. A final note of appreciation is given to the University of North Texas where I had the privilege of completing coursework with highly qualified and knowledgeable faculty. Additionally, appreciation is given to the University of North Texas for the vision and courage to offer the distance learning doctoral program. A special note of recognition is given to the supportive members of Cohort 1, Jan Bennett, Brenda Baker, Shelley Sweat, and Johney Chandler.

A personal note of appreciation is given to my two loving daughters, Leah and Kate Ridinger for their encouragement. Also, this final project could not have been completed without the support of my parents, Bill and Vale Holland. A note of gratitude is also given to my sister, Paula Swenson, for her assistance in editing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

Introduction

"Never before in history has there been such a demand for American educators to 'stand and deliver' regarding their performance" (Streshly & Newcomer, 1994, p. 62). Despite 20 years of school reform efforts, the public still perceives American schools as failing. Everyone from politicians to business leaders is calling for more accountability, more testing, and more involvement in schools (Davies, 1997). Local, state, and federal agencies share in the highly political supervision and review of public schools. The extent to which each agency is responsible for the quality of education is a great concern. Although historically the major public issue was school finance (Ramirez, 1992), the public's right to know has expanded, and as a result, the educational system is held responsible for student success. This public scrutiny began with "A Nation at Risk" (1983) and continued with approximately 300 governmental reports, commission papers, and educational task forces (Cuban, 1992; Decker, 1988; San Miguel, 1996; Rathbone, 1988).

Because of the publicity, national concern has mounted over the ability of public schools to provide students with quality instruction. Streshly and Newcomer (1994, p. 62) wrote, "America has been convinced that its school system is a failure, and has determined to reverse its fortune." Brown (1990) stated that educational accountability endures as a "slippery concept" in which few believe that the outcomes of education are under anyone's control (p. 2). Despite the widespread consensus that schools must be

accountable for student performance, few agreements exist on standards for accountability (Harrington-Lueker, 1998; Webb, 1995). The question of what constitutes knowledge and how it can be demonstrated often remains unanswered. Although logical in theory, the issue of measuring knowledge becomes controversial in practice. Analysts often speak of "litmus test" issues as those that delineate and separate political forces (Theobald, 1995, p. 462). For education, accountability is the new litmus test that most clearly separates the views of those with power to demand accountability and those in education to whom the demands are made.

Regardless of increased accountability systems, the public perceives that students are not learning enough for the world that awaits them (Banks, 1994; Davies, 1997). The future of students at the turn of the century depends on brainpower rather than horsepower (Anrig, 1992). Responding to this reality, some policymakers promote reform by posing test-based accountability standards on students, teachers, and schools. The use of collected and reported data is the critical issue. The resulting push for accountability has led to a rush to test in the belief that the two are synonymous (Anrig, 1992; Banks, 1994; Clinchy, 2001; DeBray, 1999; Fetler, 1990; & White, 1994). The result is that the government is spending valuable tax dollars and teachers are devoting valuable instruction time to test preparation while students continues to be inundated with testing (Oldham, 1993).

As educators deal more effectively with accountability, perhaps greater issues will begin to occupy the public's attention. Unfortunately, until the cry for accountability diminishes, testing and reporting continue. Anrig (1992) cautioned educators and

policymakers, "… instead of talking about testing, schools need to talk about teaching" (p. 34). According to Clinchy (2001, p. 494), "academic orthodoxy" is established when teachers are restricted to the mandated curriculum of the test. The present philosophy was simply stated by Gaines (1992, p. 17): "What gets measured gets taught. What gets reported gets taught twice as well." Or as Oldham (1993, p. 21) stated, "Measurement drives instruction."

The Texas Accountability System

The present investigation concerns itself with accountability processes and the use of comparison data in the state of Texas. Texas has an accountability system integrating accreditation ratings, student assessment, state curricula, and public reporting. The Texas Education Agency has established this integrated system through collaboration with educational professionals, professional organizations, and policymakers. The accountability system is based on eight guiding principles to ensure equitable and fair treatment across the state. As enumerated in the state accountability manual, the guiding principles are: student performance, recognition of diversity, system stability, system compliance, appropriate consequences, local program flexibility, local responsibility, and the public's right to know (Accountability Manual, 2000).

When the term integrated is used to describe the Texas system, the use is well supported. The accountability system is a composite of the state assessment, curriculum, accreditation, and professional development. The statewide testing program, Texas Assessment of Academic Skills (TAAS), serves as the state performance standard or

measurement used in assessing the implementation of the Texas Essential Knowledge and Skills (TEKS). As changes are made in the required curriculum (TEKS), adjustments are made in the TAAS testing system. Student performance also impacts the teacher appraisal system (Professional Development and Appraisal System) by awarding the teacher credit for the campus rating. For teachers on low performing campuses, points are added to the appraisal record if the campus is ranked in Quartile 1 or 2 by the campus comparable improvement report. If the campus is ranked in Quartile 4, no points are added to the teacher score. Additionally, principal appraisal and superintendent appraisal include student performance domains based on TAAS performance. The accountability for student performance does not stop at the campus or district level. The state of Texas is divided into 20 regions, each with an education service center that serves as a liaison between the state and local education agencies and that provides professional development and technical assistance. Each education service center is held accountable for student performance. An education service center executive director's appraisal also includes domains based upon TAAS performance, attendance, and the dropout rate for all the students served in the region (Accountability, 2000).

The Texas system also utilizes a reward and sanction component initiated by the accreditation rating system. Base indicators determine the accreditation ratings for districts and campuses. The accountability system has three base indicators: TAAS performance in math, reading, and writing; attendance rate; and dropout rate. With the enormous importance of the accountability ratings and reporting systems, the TEA, with input from focus groups, defined and developed a system of procedural safeguards. The

inclusion of such safeguards provides a system that is both effective and equitable. One major procedural safeguard is the inclusion of only those students within the district as of the last week of October of the particular testing year for accountability purposes. With this safeguard, schools are only accountable for students actually served within the district for a reasonable length of instructional time. Another procedural safeguard considers small numbers in assigning accreditation ratings. For each accountability indicator, a minimum number of students is required before assigning campus and district ratings, comparable improvement growth scores, or special acknowledgements. In addition, system safeguards, including data quality audits and special analysis reviews, ensure that campus and district ratings are based on properly reported and measured data (Accountability, 2000).

One major component of the state accountability system, the Academic Excellence Indicator System (AEIS), serves as the basis for reporting campus rating, student performance, and campus demographics to the public. At the core of the system is the campus report, the AEIS, which the Texas Education Agency provides annually. This AEIS includes the state-assigned accreditation ratings of exemplary, recognized, academically acceptable, or low performing for campuses as required by state law [TEC § 39.072 (a)]. The report contains data from student and staff demographics, special programs, and campus expenditures as compared to the state average and a comparative campus group. The Texas Education Agency provides an annual comprehensive manual as a technical resource to aid in the interpretation of data and the effective use of the report. The AEIS report as required by statute, TEC § 39.054, must be utilized in the

campus site based decision-making process and be reviewed publicly each year. Therefore, the reporting of data becomes very important in decision-making, planning, and campus improvement (Texas Education Code, 1999).

Because of the accreditation ratings, the most controversial information on the AEIS report is the TAAS results that are reported by student group and subject. Besides TAAS information, the AEIS provides an extensive data source for each campus including attendance, dropout rate, budget, and demographics by student, faculty, and program. The statewide Public Education Information Management System (PEIMS) collects student and school demographic information. When the pieces of the state's accountability system are viewed as a whole, information collection and reporting are common threads running through the state's educational infrastructure.

As a part of AEIS, one statistic, called comparable improvement, attempts to establish like campus groupings for the purpose of comparing improvement gains in student performance as required by statute (Accountability, 2000). Comparable improvement is defined in the accountability glossary as "a measure that shows how student performance on the TAAS test has changed (or grown) from one year to the next, and then compares that growth to that of the 40 schools that are demographically most similar to the target school" (Accountabililty, 2000, p. 6). TEA constructs a unique comparison group of 40 campuses using six demographic characteristics in order of dominance (Accountability Manual, 2000, p. 47). The statute directs that the grouping is based on likeness as determined by the percentage of low income, white, Hispanic, African American, and limited English proficient students in the order of dominance for

each campus. Although not required by statute, the TEA added the percentage of mobile students, who are present less than 83% of the school year, to the selection process for the campus cohort. The comparable improvement analysis process recreates the comparison group each year to ensure a timely match by campus type and demographics.

The order of dominance is the key component to constructing the 40-campus comparison grouping for comparable improvement. First, all campuses are grouped by campus type (elementary, middle, high, or multi-level). In the comparison for an elementary campus, for example, the original 100 campuses would include only elementary campuses as potential selections. The six characteristics for the target campus are then ranked from highest to lowest percent. Next, using the most dominant characteristic of the six characteristics, 100 campuses are chosen from the campus type selection pool that are the most similar to the target campus. From this point, the process changes from selecting campuses to removing campuses from the pool. The next most dominant characteristic is used to eliminate 10 campuses of the 100 that are the most dissimilar, in either direction. The process continues until only 50 comparison campuses remain. Finally, the group is reduced to 40 by removing the 10 campuses with the most dissimilar of the less predominant characteristics of the accountability student-group characteristics -African American, Hispanic, white, and economically disadvantaged. Because only student groups that are included in the accountability system are used, the percentage of limited English proficient and mobility are not included in the final reduction from 50 to 40 campuses.

Comparable Improvement Quartile Rankings

Once the comparison group is established, comparable improvement includes the campus's designation into a quartile based on the average growth on TAAS in reading and mathematics over a multi-year period. The 40 campuses are ranked according to their average growth. Using the TAAS Texas Learning Index (TLI) that reflects actual growth from year to year, the ten campuses with the highest average TLI are assigned Quartile 1 placement for reading and math. Ten campuses are placed into Quartile 2 and into Quartile 3 as the ranking continues. The ten campuses in Quartile 4 reflect the lowest average growth in the TLI of the campus group.

In addition to publicly reporting comparable improvement in the AEIS reports and the School Report Card, campus quartile placement impacts both rewards and sanctions in the accountability system. As a recognition award, a campus placed in Quartile 1 with 50% of the matched students scoring above 84 TLI, called high performing students, receives a comparable improvement acknowledgement on the campus AEIS report for high performance in reading or math (2000 Accountability Manual). As a monetary award, the Texas Successful Schools Awards System (TSSAS) rewards campuses for demonstrating significant gains as measured by comparable improvement in reading and math. Each spring, approximately 800 qualifying campuses that are rated acceptable and above, ranked Quartile 1 for reading and math, and have moderate exemption rates for each subject area tested on TAAS receive TSSAS awards. TEA establishes the baseline for limited English proficient and special education exemption rates each year. For the 2001 TSSAS award selection, the acceptable

exemption rates were at or below 14.5% for reading, 13.6% for math, and 15.6% for writing. The amount of financial award is equal to $7.20 multiplied by the average daily enrollment, and awards range from $500 to $5000 per campus. The principal and the campus improvement team must determine how the TSSAS award funds are spent. Receiving this award also entitles the campus to apply for $500,000 of additional funds for increased parent involvement activities through a competitive grant process (TEA website, 2000).

As a sanction, the negative publicity from scoring in Quartile 4, the lowest quartile, has additional implications for campus staff. Leading a campus ranked as Quartile 4 can impact principal assessment with average TLI growth and quartile placement serving as two student performance measures. As perceived by the public, the term like campuses is comparable to equal campuses. Therefore, the public views the campuses within the comparable improvement grouping as facing equal challenges, barriers, and expectations.

Comparable Improvement Student Characteristics

Poverty

Although the current method of sorting appears fair, do inherent inequalities impact campus rankings? How comparable are the campuses within the forty-campus grouping? With the current body of research (Adams, 1994; Alexander, 1998; Berne, 1994; Brown, 1994; Coleman, 1966; Cormier, 1992; Dollinger, 1997; Fortune, 1979; Haetinger, 2000; Hemberger-Coffey, 1991; Hill, 1999; Lamberson, 1989; MacDonald,

1996; Mayer, 1997; Maume, 1998; Perkins, 1992; Pierre, 1994; Plecki, 1991; Thompson, 1996; Smith, 1999; & Yong, 1987) on the negative impact of poverty on student performance, should campuses first be compared to other campuses with equal levels of poverty? The question of comparison based on poverty becomes more relevant as America becomes more impoverished (Berne, 1994, Council of Economic Advisors for the President's Initiative on Race [CEAPIR], 1998; Mayer, 1997; Reigel, 1992; Webster, 2001; & Welch, 1995). In an article in the Wall Street Journal, Hunt (1999, p.19) wrote of the increasing inequality where the "kids who are living in poverty are poorer than ever." Clinchy (2001, p. 495) stated that the "always unlevel playing field is growing even steeper for the poor and for minorities as each day passes." How does this increasing poverty impact the classrooms and the children striving to meet higher standards? By reviewing the magnitude of the current literature, the importance of the selection process is more evident.

Coleman and his associates (1966) reviewed the relationship between socioeconomic status and student achievement, and they found poverty an important factor in a child's success in school. Following Coleman, Fortune (1979) stated that as socioeconomic status increased, achievement increased for all ethnicities. Adams (1994) found per capita income (r =. 27) and at risk classification (r =. 36) to be associated with SAT scores. Yong (1987) found median family income and Chapter 1 enrollment were the best predictors for the ACT test. Alexander (1998) found that the students in poor schools enrolled in 35% more remedial classes and 7.1 fewer core classes per week than did students in more affluent schools. Hill (1999) found that the level of achievement

(reading r = .17 and math r = .27) was directly tied to the economic level of the neighborhood where the students attended school. Brown (1994) found free lunch status predictive for reading levels, and MacDonald (1996) found socioeconomic status predictive for both reading (r = .72) and math achievement (r = .57). Cormier (1992) found that two socioeconomic indicators, parent education level and community occupation index, explained the variances (r= .63) in student math achievement. Hemberger-Coffey (1991) concluded that socioeconomic variables including district wealth and occupational status of the adult population constituted strong correlations to student performance. Lamerson (1989) found low-income enrollment to be a chief predictor for quartile placement on achievement tests for 10^{th} graders. Newman (1998) stated that the primary factor affecting success in school was poverty. Pierre (1994) found that family poverty was more predictive of low student success for boys than for girls. Although Plecki (1991) found that poverty impacted performance differently at various school sizes, socioeconomic status was the strongest predictor for student success. Thompson (1996) reported that school and family socioeconomic status was strongly correlated (r = .82) to total achievement levels. Dollinger (1997) found that poverty impacted student attendance; Maume (1998) found that a low socioeconomic level of the community influenced juvenile delinquency.

Research studies from Coleman (1966) to the present continue to show that poverty does impact student achievement. The statement that, “Poverty knows no color” (Connerly, 1999, p. 42) perhaps complicates the interpretation of many of the findings. In summary, although each student characteristic may impact achievement, poverty appears

to be one of the most pervasive obstacles to success in school. In determining the comparable improvement comparison group, the characteristic used to select the initial 100-campus group has been based on the most dominant of the six characteristics, not the most predictive of student performance. When campuses are compared to other campuses for accountability, what student characteristics are the most dominant or the most significant? If comparable improvement compares like campuses for average growth, would basing the comparison on comparable poverty levels be the most informative and beneficial method?

The Problem

The problem for this dissertation is to determine how comparable is comparable improvement for campuses in Texas. In the original study, the Texas Education Agency (Haetinger, 2000) conducted a canonical analysis of the relationship of the six characteristics utilized in comparable improvement to student performance. In this analysis, student performance was measured by passing all TAAS tests at the grade level. According to the 1995 analysis, the strongest predictor at the elementary level for student performance was economically disadvantaged classification with a correlation coefficient of $r = .72$. The other comparable improvement characteristics also were found to have a moderate correlation to student performance although not as strong as poverty: African American with $r = .38$, Hispanic with $r = .41$, limited English proficient with $r = .32$, and mobility with $r = .46$. The percent of white students was the least predictive, $r = -.64$, for student performance as measured by passing all test taken. The comparable improvement

characteristics, except for percent of white students, individually had a moderate to strong correlation to student performance, but no additional studies have been conducted to discern the relationship of the six characteristics to TLI average growth, award eligibility, quartile placement, or campus grouping process.

Purpose

The purpose of this study was to examine an alternative strategy for determining comparable improvement. The percent of economically disadvantaged students was used to select the initial pool of comparison campuses. Based on the literature on student performance, this variable may be the most salient variable for comparison purposes. After using this alternative strategy, the impact on quartile placement, comparison group TLI average growth, award eligibility, and changes in campus grouping between the current and the alternative comparable improvement process were examined.

Research Questions

1. What are the differences in quartile placement based on the current order of dominance process when compared to an alternative process using the percentage of poverty as the initial selection characteristic?

2. What is the difference in the campus average TLI growth and group average TLI growth based on the current order of dominance process when compared to an alternative process using the percentage of poverty as the initial selection characteristic?

3. What changes occur in the award eligibility when the current order of dominance process is compared to an alternative process using the percentage of poverty as the initial selection characteristic?

4. What changes occur in campus comparison group composition when the current order of dominance process is compared to an alternative process using the percentage of poverty as the initial selection characteristic?

Limitations

Limiting factors for this study include the following:

1. The data used in the AEIS demographics include self-reported data by the district as taken from the Public Education Information Management System (PEIMS). This may influence the accuracy of the campus demographics included in the analysis.

2. The measure for student performance is the state assessment, Texas Assessment of Academic Skills, as administered following the state test administration guidelines. The results will be not applicable to other testing systems or other test administrations and are dependent on the dependent on the reliability and validity of the TAAS scores.

3. The comparable improvement analysis is based on “matched” students in determining average TLI growth over a two-year period. Matched students are students at each campus whose TAAS results can be found in the current and prior test anywhere in the

state. When students are not matched, no TLI growth can be computed, and these students are not included in the comparable improvement analysis.

Delimitations

1. The analysis utilizes the statewide accountability report system, AEIS, in the ranking of comparable improvement. The results will not be applicable to other accountability systems or public reports.
2. Only elementary campuses are included in the analysis; therefore the findings do not apply to junior high or high school campuses.
3. Elementary campuses newly opened are not assigned an accreditation rating, and comparable improvement is not calculated. Campuses without an accreditation were not included in this study.

Definition of Terms

1. Academic Excellence Indicator System (AEIS): the Texas public reporting procedure that includes all indicators used in the Texas accreditation system and student, staff, and budget profiles.
2. Accountability: a system of commitments, policies, and practices to assign responsibility for student performance to local educators.
3. Accreditation rating: the status according to the campus performance for campuses in Texas on the identified indicators: exemplary, recognized, acceptable, and low performing.

4. Base indicators: the indicators used for determining Texas accountability ratings, including college admissions testing results, Texas Assessment of Academic Skills (TAAS) /Texas Academic Skills Program (TASP) equivalency, recommended high school participation, and comparable improvement for reading and mathematics.

5. Comparable improvement: a measure that shows how student performance on the TAAS test has changed (or grown) from one year to the next and then compares that growth to that of the 40 schools that are demographically most similar to the target school for campuses in Texas.

6. High-performing students: students scoring at or above a TLI score of 85. These high-performing students are excluded from the campus average growth calculation for campuses in Texas.

7. Low-performing students: students scoring at the minimum score in either the current or prior test administration.

8. Order of dominance: the process used in the Texas accountability system to determine the unique 40-campus comparison group used in comparable improvement. The order of dominance is determined by ranking the campus's six demographic characteristics from highest to lowest percentage.

9. Professional Development and Appraisal System (PDAS): The Texas teacher appraisal process used annually to appraise all instructional staff.

10. Percent African American: the number of African American students enrolled multiplied, by 100, and divided by the number of students in campus membership.

11. Percent Hispanic: the number of Hispanic students enrolled, multiplied by 100, and divided by the number of students in campus membership.

12. Percent white: the number of white students enrolled, multiplied by 100, and divided by the number of students in campus membership.

13. Percent economically disadvantaged: the number of economically disadvantaged students enrolled, multiplied by 100, and divided by the number of students in campus membership.

14. Percent limited English proficient: the number of limited English proficient students enrolled multiplied by 100 and divided by the number of students in campus membership.

15. Percent mobile: the number of students in campus membership less than 83% of days taught multiplied by 100 and divided by the number of students in campus membership.

16. Practical significance: the level of significance reflecting a considerable relationship beyond statistically significance due to sample size.

17. Public Education Information Management System (PEIMS): the data collection and reporting process required of all districts in Texas that includes student, teacher, budget, and program data.

18. Reward: recognition or monetary incentives for high performing or improving schools.

19. Sanctions: punitive actions against low performing schools.

20. School Report Card (SRC): a public reporting procedure highlighting a district's performance on the state performance indicators.

21. Site Based Decision Making (SBDM): the process for collaborative planning and decision making as required for all campuses by Texas Education Code (TEC) §11.252.

22. Statistical significance: when the probability that the result occurred by chance is less than the predetermined alpha level.

23. Texas Assessment of Academic Skills (TAAS): the state of Texas assessments for grades 3 through 8 and grade 10 in reading, writing, mathematics, social studies, and science.

24. Texas Education Agency (TEA): the state education agency for Texas.

25. Texas Learning Index (TLI): a score on the state assessment that reflects actual growth in one subject area from year to year on the TAAS.

26. Texas Successful Schools Awards System: the state monetary reward system for schools with the highest sustained success or the greatest improvement in the Texas accountability indicators.

Significance of this Study

This dissertation will determine how reliable a poverty-driven alternative strategy for determining comparable improvement is comparable to the current method of determining the group TLI average growth, award eligibility, quartile placement, or campus grouping process. If this study finds that campus comparisons of student performance is, in fact, significantly impacted by poverty, it would be helpful to high poverty schools in presenting a case for a change of public policy regarding the state's recognition and reward system. Possibly, the determination of rewards based on quartile

placement as presently configured will prove to be inappropriate. This information could potentially alter the reward process, thus establishing a more equitable comparison for high poverty campuses. Importantly, this alternative comparison process still allows for a value-added evaluation of struggling schools. The alternative process changes the reference or comparison group and bases comparable schools on a potentially more salient criterion, poverty. The results of this study could provide state officials with empirical evidence for policy decisions regarding the use of comparable improvement for special acknowledgements, public recognition, and monetary rewards.

Comparable improvement was first created in statute in 1993, but implementation did not begin until the 1995-96 school year (Accountability Manual, 2000). The calculation and application of comparable improvement are assigned to the commissioner of education. Furthermore, although state statute defines five of the six comparable improvement characteristics, the Texas Education Agency does have control over the process for sorting and compiling like campuses. Information on the impact of poverty could prove invaluable to TEA. Because of the periodic review of the state accountability system, this study is timely and could provide quantitative data for decisions at the state level.

Last, this study could furnish additional insight into the relationship of TAAS performance, poverty, and comparable improvement. Because the AEIS report is used in campus decision-making, it is important for all site-based teams to understand the relationship between comparable improvement and student performance, particularly as decisions are made to better serve students.

Conclusion

A high quality accountability system incorporates public reporting with sound judgment and robust procedural safeguards in the search for improved student performance. In a state as diverse as Texas, the use of comparable improvement in determining accountability recognitions and financial rewards must be based on equitable and fair comparisons. The alternative comparison process suggested here might be useful in providing a more equitable process. The present study investigated the equality, or lack thereof, between the current process and the proposed process. However the possibility remained that other variables, other than those used currently by TEA, predict student performance better.

For Texas schools, “comparable” should actually mean “comparable.” Although the purpose of any accountability system is improved student performance, the public pressure placed on educators by the state accountability system is often negative and potentially harmful (Clinchy, 2001). In the quest for improved student performance, the state must be passionate in the pursuit of fairness and equity for all campuses. The present investigation evaluates a potentially more equitable process for defining campus comparison groups.

CHAPTER II

Background and Review of Literature

Introduction

The purpose of this study was to examine the impact of poverty on quartile placement, comparison group Texas Learning Index (TLI) average growth, award eligibility, and changes in campus grouping between the current and the alternative comparable improvement process. Included in this discussion is a summary of literature and research on accountability systems, specifically in regard to educational indicators, reporting procedures, and current research findings. Additionally, to better understand the relationship of poverty to student achievement, research on the impact of student demographics such as poverty, mobility, ethnicity, and language proficiency on student achievement are also to be reviewed. With the statement by Darling-Hammond (1992) that the mere reporting of educational indicators can change teacher behaviors, understanding the relationship between poverty and student achievement is foundational in determining the fair and equitable use of comparable improvement reporting.

Poverty

During the past thirty years, an increasing amount of research has been devoted to examining the impact of poverty on student achievement, causing educators alarm. The current body of research (Adams, 1994; Alexander, 1998; Brown, 1994; Coleman, 1966; Collier, 1994; Cormier, 1992; Dollinger, 1997; Franklin, 1992; Fortune, 1979; Haetinger, 2000; Hemberger-Coffey, 1991; Hill, 1999; Lamberson, 1989; MacDonald, 1996;

Maume, 1998; Newman, 1998; Perkins, 1992; Pierre, 1994; Plecki, 1991; Thompson, 1996; Smith, 1999; Soltz, 1991; & Yong, 1987) found that poverty has a strong negative impact on student performance. Although poverty was defined differently in the various research studies, for example, free-lunch status, median family income, occupational index, or community income status, the connection between a student's poverty level and a student's achievement was a common finding.

In a controversial study in student achievement, Coleman and his associates (1966) reported on educational equality in America. The Coleman study reviewed the achievement of more than 500,000 students and found a strong relationship between socioeconomic status and student achievement. If poverty was the single most important factor in a child's success in school as Coleman found, educators and policymakers had difficult issues to face. The study, while forcing the nation to look at both the equity and equality of American schools, served as an impetus for many subsequent research projects.

In a much smaller study, Fortune (1979) found that as the student's socioeconomic status increased, the student's achievement increased. This relationship was found similarly across all ethnicities. Although the Fortune study discovered that ethnicity greatly impacted student achievement, poverty served as a key factor in a student's success in school. In the past decade, studies continued to reveal that a high poverty status was associated with low student test scores. After reviewing Stanford Achievement Test (SAT) results, Adams (1994) found that low per capita income levels were associated with low SAT test scores. Using the three at-risk factors of percentage below poverty level families, percentage of children not living with parents, and

percentage of adults without high school diplomas, a high at-risk index was also associated with low SAT scores. Overall conclusions were that as per capita income (r = .27) increased and the risk factors (r = .36) decreased, the SAT scores increased for the students.

Mayer (1997) reviewed the relationship between verbal ability and student achievement as measured by the Peabody Picture Vocabulary Test (PPVT) and the Peabody Individual Achievement Tests (PIAT) and poverty. Quintiles of socioeconomic levels were determined based on family income. On the PPVT, the poorest quintile group had a mean score 13.5 points below the most affluent group. For the PIAT math test, a similar gap existed. The poorest quintile mean score was 96.9 while the most affluent quintile mean score was 104.4. In PIAT reading, the poorest quintile's mean of 101.9 was 8.3 points below the most affluent quintile's mean of 108.2. In addition to verbal ability and achievement, Mayer included an analysis of the Behavior Problems Index (BPI). With the BPI, a high score indicates more discipline and behavior problems. Through an interview with the student's mother, information was gathered concerning the child's behavior. Mayer reported that the poorest group's mean score of 109.7 was 6 points higher than the most affluent group's mean of 103.7. The BPI score increased as the poverty levels increased across the quintiles. The poorest quintile also had the highest percentage of girls becoming teenage mothers. Additionally, 40% of the girls in the poorest quintile became teenage mothers while only 4.9% of the top income quintile became teenage mothers. Mayer (1997, p. 53) concluded, "Children from low-income families score lower on tests of cognitive skill than children from affluent families, are more likely to have babies as teenagers or become

young single mothers, and are more likely to drop out of high school and receive fewer years of education."

Berne (1994) reported similar findings in a review of the New York Pupil Evaluation Program (PEP) and poverty status. The New York PEP assessment was administered to all public and private school students as a measure of academic achievement. The state developed reference point, passing score, represented minimum achievement. Of the percent of the students passing the 1993 PEP assessment, a 36 point gap was found between the campuses with the lowest poverty, 0-20% poverty, and the campuses with the highest poverty, 81-100% poverty. The achievement gap between the poorest and richest schools was consistent across all district types except for rural districts. For only rural districts, students in the 61-80% poverty quintile had a 98% passing rate while the students from the 0-20% poverty quintile had a 91% passing rate. The high poverty school scored 7 points higher than the low poverty schools. Regarding the overall results, Berne (1994, p. 13) wrote, "The highest poverty schools are consistently among the lowest scoring schools, and the differences between the high and low poverty schools are substantial. This is a clear indication of the strength of the outcome inequalities in the state."

Thompson (1996) studied the impact of socioeconomic status on total achievement levels. He found a strong correlation ($r = .82$) between school and family socioeconomic status and the achievement of students. In another study on achievement levels, Cormier (1992) reviewed school factors and community economic indicators to predict student achievement scores on the Maine Educational Assessment (MEA). Cormier found that the only consistent trend in all the student performance measures was the socioeconomic status ($r = .63$) variables which, unlike the teacher inservice and other school variables, are all outside the direct control of

the school district. With the findings, Cormier (1992, p. 148) advised states to find alternative assessments "less influenced by socioeconomic status variables" in designing school reform models.

In a similar study, Hemberger-Coffey (1991) evaluated the Connecticut Mastery Test performance and cumulative district averages for eighth grade students in relation to three demographic and nine financial input variables. All three of the socioeconomic variables of district wealth ranking, educational background, and occupational status of the district's adult population displayed strong positive correlations to student achievement, again finding that as income increased so did the achievement levels. The highest positive correlation of $r = .67$ was found for occupational status and $r = .63$ for educational background, a close second. In the ranking of all 12 variables, the socioeconomic variables represented the three highest correlations to student achievement. In regard to subject area, the effect of each socioeconomic variable was more pronounced in mathematics than in language arts.

Using the Illinois school report card data, Lamberson (1989) examined the relationship between the student and school variables and 10^{th} grade student performance in reading. The student variables were student attendance rate, student mobility rate, and percent of poverty. School variables included average class size, pupil-student ratio, operating expenditures for instruction, and pupil-administrator ratio. In reading, school variables were not statistically significant. However, a set of three variables, student attendance rate, proportion of low income, and district enrollment, was statistically significant ($r = .30$). The proportion of low-income enrollment was negatively related to student achievement, while the other two variables were positively related to achievement. Of the three student variables, Lamberson found the

percentage of low-income enrollment and student attendance the most prominent. Similarly to Cormier and Fortune, Lamerson found that as the percentage of low-income enrollment increased, the percentage of students scoring in the top quartile decreased.

In another Illinois study, Yong (1987) studied the impact of district wealth and size upon student and school performance. Using selected accountability indicators from the Illinois report card, Yong found median family income was a good predictor for student and school achievement. The wealthier schools generally reported higher American College Testing (ACT) scores, attendance rates, and higher graduation rates, as well as lower pupil-teacher ratios, higher salaries, and per pupil expenditures. Poor schools had lower ACT scores and lower graduation rates. In regard to predicting school performance, the percent of Chapter 1 students was found to the best predictor. Specifically, a high school with 60 percent Chapter 1 enrollment had an ACT composite score of 13.1, while a high school with 2 percent Chapter 1 enrollment had a 20.6 composite ACT score. Yong (1987, p. 120) stated, “Poverty has been shown to have a depressing effect on student performance.”

Using reporting indicators, Louisiana’s progress profiles for students in 1,336 districts, Franklin and Crone (1992) reviewed the relationship between school and student characteristics. Again, poverty surfaced, along with suspension and dropout rates, as inversely related to student achievement. Using ANOVA procedures with class size, teacher certification, student dropout, and student attendance, the findings revealed little impact on student test scores by school size, school type, or class size. For high poverty populations, large urban schools had an increased negative relationship to student achievement. Franklin and Crone (1992) concluded

socioeconomic status was more predictive of student success than either school size or attendance.

In a study by Smith (1999), accountability input variables were analyzed to determine if district spending was associated with higher student achievement when student demographics were controlled. The study included 64,000 students in Minnesota public schools and their performance in mathematics and reading. All student-level background variables including disability condition, poverty, limited English proficiency, and ethnicity were associated with decreased student achievement in reading and math. Not surprisingly, free lunch eligibility was found a stronger predictor than reduced lunch eligibility. Although all student variables impacted student performance, Smith (1999, p. 176) concluded, "Without a doubt, poverty, as indicated by students' eligibility for free or reduced-price lunch, is strongly associated with lowered achievement."

Plecki (1991) explored the relationship between school size, student achievement, and poverty. Using 4,337 California elementary students, Plecki found that socioeconomic status explained the variance in student performance more than school size. Similarly to Franklin and Crone, findings on the relationship between school size and student performance were mixed. Once again for high poverty populations, large urban schools had an increased negative relationship to student performance. The impact of poverty was consistent across each school enrollment size.

Hemberger-Coffey (1991) reviewed the impact of multiple student and district characteristics on reading and mathematics achievement. For reading, the three highest correlated variables were adult occupational status ($r = .63$), educational background of the

adults in the district (r = .60), and district wealth as measured by property tax and per capita income (r = .56). Similar results were found for mathematics, occupation status (r = .67), educational background (r = .63), and district wealth (r = .58) were related to student achievement levels. MacDonald (1996) analyzed the relationship between student performance and demographic variables for 54 recognized districts in Alaska. The strongest predictor for student performance was found to be student socioeconomic status (reading r = -.72, math r = -.57). As found by Cormier (1992), Fortune (1979), and Lamerson (1989), MacDonald found that as the percentage of free- and reduced-lunch students increased, academic achievement decreased. Instructional expenditures were found to positively impact student achievement, while administrative and maintenance operations expenditures were found negatively correlated to student performance. MacDonald (1996, p. 107) concluded, "Student socioeconomic status was a very important predictor of student achievement in school districts in this study."

Hill (1999) examined the effects of school climate, student demographics, teacher expectations, and school-community integration on student achievement in high-poverty schools. The study included 36 high poverty campuses in a southeast urban Texas district with various student performance levels. With similar poverty levels, some campuses were rated as low performing, while others campuses were able to maintain an exemplary status. In an analysis of campus characteristics, Hill found that student background factors were the greatest predictors for student achievement (math r = .34 and reading r = .37), followed by teacher expectation and climate. The instructional program factors produced the least impact (math r = .20 and reading r = .21) on student performance. Hill (1999, p. 94) stated, "In general, it appears that the student's background (what she/he brings to school) definitely impacts student achievement."

Alexander (1998) examined the relationship between the state mandated curriculum, student participation, and student achievement in New York public high schools. A six-year review was conducted utilizing report card and interview data focusing on the impact of policy upon student achievement. The implementation of higher standards and a more rigorous curriculum had been expected to close the achievement gap; however the performance of minority and poor students remained below their more affluent, white peers. Following the implementation of increased standards policy, Alexander (1998, p. 122) concluded, "a student's socio-demographic background and where he/she attends school may have an adverse effect on the quality of the curriculum he/she receives." In addition to finding fewer academic core classes or advanced courses in high poverty, high minority schools, Alexander also found lower passing rates. Alexander (1998) described the impact of high poverty schools: "…it appears that the quality of education to which students are exposed, as well as their subsequent achievement, is tied to where they attend school."

In a qualitative case study, Dollinger (1997) utilized interview and focus group formats to investigated the factors impacting absenteeism, particularly excessive absenteeism, for high school students. Among the family and personal characteristics impacting absenteeism, poverty was a primary factor. In a similar study on juvenile delinquency, Maume (1998) found that the socioeconomic level of the community impacted delinquency and student commitment to succeed.

Although not the only factor, the research studies found that poverty impacts student achievement, test scores, passing rates, achievement levels, attendance, and juvenile

delinquency. In general, poverty was one of the more consistent predictors for student success in school.

Poverty and the Other Comparable Improvement Characteristics

For comparable improvement, the minority status is limited to African American, Hispanic, and white and does not include Native American or Asian. In the research studies reviewed, ethnicity was more inclusive than comparable improvement. Although most of the research studies focused upon either ethnicity or poverty, many studies (Alexander, 1998; Bourke, 1998; Collier, 1994; Edington, 1984; Fortune, 1979; Holman, 1993; McClure, 1999; Smith, 1999; & Whitus, 1996) have analyzed the impact of multiple student characteristics upon student achievement. For example, Brown (1994) found ethnicity and mobility to coexist with poverty as significant predictors of student performance. Alexander (1998), Edington (1984), Holman (1993), and Smith (1999) found that student characteristics of poverty and ethnicity had a strong negative impact on student success in school. Haetinger (2000) found that poverty, ethnicity, mobility, and language proficiency were all highly indicative of student success, but poverty was the highest predictor.

When socioeconomic status is excluded, the research findings on the impact of ethnicity and its relationship to student performance are less consistent. Allen (1991), England (1993), Jencks (1998), Landry (1997), Shepherd (1998), and Zankofski (1999) found that ethnicity did impact student achievement to some degree. Allen (1991) found

statistically significant differences in the mean scores of whites, African Americans, Hispanics, and Asians on reading tests. Landry (1997) concluded that ethnicity, along with gender, was more likely to be an at-risk indicator for not completing baccalaureate nursing programs. Shepherd (1998) in researching magnet schools found that gender and ethnicity were the main factors in science achievement for males and white students. In researching community college success, Zankofski (1999) found ethnicity had an effect on math performance by course grade and passing rate.

In The Black-White Test Score Gap, Jencks and Phillip (1998, p. 1) state "African Americans currently score lower than European Americans on vocabulary, reading, and math." For the 1989 college cohort, the average SAT score for African American students was 1131 and for white students was 1306. Even where the students' SAT scores were controlled, African American students had a grade point average (GPA) .33 points below their white peers. Additionally, only 19% of the African American students earned a GPA above the mean for their white peers. Although the gap appears before kindergarten, it continues to adulthood. According to Vars and Bowens (Jencks & Phillips, 1998, p. 466), the achievement gap between African American and white students is both "substantively significant, as well as statistically so." Although income inequality between the African American and white population may impact the test score gap, Jencks and Phillips (1998, p. 9) concluded, "It is quite small."

Although poverty was the primary issue in the study, Berne (1994) included a comparison of student achievement and ethnicity using the Pupil Evaluation Program (PEP) assessment. Berne found that the low minority districts scored higher than the high

minority districts. Districts with 0-20% minority enrollment had a 93% passing rate on the PEP assessment. The districts with 81-100% minority enrollment had a 58% passing rate. Again as in the Berne's poverty study, the high-minority, rural districts had a higher passing rate than the low-minority rural districts. The districts with 0-20% minority enrollment had a 91% passing rate while the districts with 81-100% minority enrollment had a 98% passing rate, a 7-point difference. Berne (1994, p. 13-14) wrote, "Although the relationships are not uniform for every district type, they are strong and consistent."

Conversely, many studies (Bonsangue, 1992; Brown, 1998; Eineman, 1992; Kaminski, 1999; Nine, 1990; & Seidenfeld, 1991) have found that ethnicity was not a statistically significant factor. At the college level, Bonsangue (1992) stated that achievement among minority students in mathematics and science was associated more with lack of experiences than ability. In a similar conclusion, Brown (1998) found that after participating in a five-week extended year program, minority students made significant gains: African American students experiencing higher gains than white students. Eineman (1992), while studying the impact of retention, stated that mental aptitude (r .39) was a more statistically significant factor to student achievement than ethnicity (r .33). Nine (1990) found that motivation and attendance, not poverty or ethnicity, impacted performance. Seidenfeld (1991) researched the effect of instructional strategies in the area of spelling and found that teaching methodology, not ethnicity, contributed to word retention.

In regard to the remaining comparable improvement characteristics, Brown (1994), Kaminski (1999), and Newman (1998) examined mobility and student achievement. In a study to

determine the impact of mobility on achievement levels, Brown (1994) reviewed the performance of 434 primary students in the San Bernardino Unified School District. The researcher found mobility negatively impacted student performance, particularly the youngest students in first and second grade. In regard to gender and ethnicity, Hispanics and girls were the most affected, and five or more moves were significantly impacted student performance. In addition to mobility, the study also researched the impact of poverty on student achievement. Brown (1994, p. 99) found, "The poorer a student, the more likely he/she had lower reading performance."

Although each child's situation was unique, poverty and mobility provided additional constraints to the school success of the homeless (Newman, 1998). Through interview, observations, and document reviews, Newman examined the complexity of homelessness in Southern California. Among the factors impacting the homeless child's ability to succeed in school were devastating poverty, personal and social problems, high mobility, and lack of family support for schooling. The homeless child faced multiple barriers in order to be successful in learning.

Limited English proficiency, the last comparable improvement characteristic, affects multiple minorities from Hispanic, Laotian, Cambodian, German, to Japanese. Ranking second only to California, more than 10% of all students in Texas or reportedly 555,470 students are identified as limited English proficient (LEP) according to the Public Education Information Management System (PEIMS) reporting for 2000 (TEA, LEP Study, 2000). Although 91.6 % of the LEP students or 508,907 students are Hispanic, significant numbers of students speak other languages. Of the remaining 15,098 students, Vietnamese represent the largest language group,

followed by Chinese and Korean (LEP Study, 2000). From 1992 to 1998, there was a 44% increase in the number of LEP students served in Texas schools (Texas Successful Schools Study, 2000). Because public school instruction is predominately in English, LEP students required the support of language instruction programs to meet the high standard as required in the state assessments. With such support systems, the research showed the negative impact of limited English proficiency on student achievement can be successfully surmounted (Texas Successful Schools Study, 2000).

Research on Successful High-Poverty, High-Minority Schools

To understand why some high-poverty campuses are successful while others are not, Texas Education Agency (TEA) has supported various research collaborations between universities and other research centers. One of the more recent collaborations, the Texas Successful Schools Study (TEA, LEP, 2000), reported that campuses implementing quality second-language programs demonstrated high success rates for limited English students. While serving high percentages of poverty, high minority, and high limited English proficient student populations, campuses included in the study reached and maintained recognized or exemplary rating as based on the Texas accountability system.

In a similar study, TEA and the RMC Research Corporation (2000) completed a study of high-minority, high-poverty middle schools experiencing continued academic success. Using data sets from TEA, campuses selected had experienced growth on TAAS

state assessment, had above 70% Latino populations, and had above 70% poverty (RMC, 2000). Characteristics found present on high performing campuses include high expectations for all students, careful disaggregation of test data, strict accountability, positive school climate, teacher expertise, and strong leadership (RMC, 2000). A national study found similar results as the Council of Chief State School Officers (USDE, 1999) reviewed the top-performing, high poverty schools. The CCSSO reviewed 1,200 successful high poverty elementary, middle, and high schools in twenty-one states. The typical campus profile had an enrollment of 509 students with 60% white, 17% African American, 13% Hispanic. All of the campuses included had 50% or higher poverty levels and scored above the average on the state assessment. The successful high poverty campuses stressed the use of state standards extensively to drive instruction, used extended time so that all students could achieve the high standards, and devoted significant funds to professional development focusing on changing instructional practice. Additionally, high accountability with strong sanctions for schools and personnel was found in 45% of the campuses.

Although the Coleman Report (1966) concluded that family background was the dominant predictor for student achievement, the effective schools research by Edmonds (Lezotte, 1986; Edmonds, 1979) identified the characteristics of successful high poverty urban schools. The effective schools research sought to understand why some high poverty schools were successful while others were not. Through this research, the characteristics that distinguished the instructionally effective from the ineffective schools

were enumerated. The characteristics became the cornerstone of the effective school research model used in school improvement efforts (Lezotte, 1986; Edmonds, 1979).

Following the effective school research, the Effective Border Schools Research and Development Initiative was a collaborative project with the University of Texas, Region 1 Service Center located in Edinburg, and several Texas border school districts (Young, 1996). The project examined the effective border school techniques to assist other campuses in improvement efforts. Schools included in this study were academically successful with predominantly Mexican-American students from poor, limited-English-proficient, non-English speaking, and/or migrant backgrounds (Young, 1996).

Other studies have found high poverty campuses experienced high success rates with high minority populations (Cavazos, 1999; Hill, 1999, McClure, 1999, Powell, 1997; Rowland, 1999; & San Miguel, 1996). Cavazos (1999) and McClure (1999) both conducted case studies of successful high minority campuses. Cavazos (1999) examined the academic success of Hispanic high school students and the principal instructional leadership skills on successful campuses. Cavazos (1999, p. 206) found that, "successful principals create a school culture of high expectations, trust, and professionalism…and require a strong emphasis on teacher accountability for student performance."

McClure (1999) focused on the successful change from low performing to exemplary status for high-minority, high-poverty campuses. Although these were high-minority, high-poverty campuses, common characteristics were found that supported the high achievement, such as strong leadership, high accountability, high standards, and an empowered staff. The successes found at high-poverty, high-minority campuses, as found

in the above studies, resulted from value-added component, such as strong leadership, shared commitment focused on campus goals, instruction centered teamwork, and high expectations.

According to the Texas 2001 student enrollment report summary (Webster, 2001), 75% of the Hispanics and 64.0% of the African American students are economically disadvantaged as compared to 20.8% of the white students. According to the Texas Successful Schools Study (TEA, 2000), one in every four students under the age of 18 and approximately 1 in 2 of African-American and Hispanic children live in poverty. As found by Welch (1995), minorities have a higher poverty rate than the general population. Reigel (1992) stated that poverty, ethnicity, and limited English proficiency are highly correlated, with positive correlations values ranging from $r = .78$ to $r = .94$. Because child poverty mirrors the poverty for racial and ethnic groups, black, Hispanic, and American Indian children have higher poverty rates than non-Hispanic white (CEAPIR, 1998). As written in the report by the Council of Economic Advisers (1998, p. 33), "Child poverty not only reflects current economic problems among poor families with children, but it is also associated with inequality of opportunity, risks to health and child development, and long-term economic disadvantage." With this redundancy among characteristics, perhaps poverty is the best spokesman for all the comparable improvement characteristics.

In summary, poverty crosses all ethnicities, all languages, and all ages. As predicted by Decker (1988), schools will have more children from poverty, more latchkey children, and children from blended families. With the increase in multi-race births and

the blurring of the race lines, it is the financial status, not skin color, which is important (Connerly, 1999). This emphasis on economic status supports Clinchy's (2001, p. 494) concept of "economic segregation" which he believes is replacing the past segregated school systems. According to Clinchy, the economically segregated schools are appearing in both the high minority inner city and the predominantly white rural areas. Unlike a clearly defined poverty status, Jencks and Phillips (1998) define ethnicity as "a social category, whose biological correlates vary geographically and historically."

Poverty, if left unchecked, continues as a significant demographic characteristic impacting student performance. In referencing the classroom, Hill (1999, p. 16) writes, "The condition of poverty is indeed a real contributor to the inequalities children bring to school and experience in school." As stated by Yong (1987. p. 120), "Thus, the education of students…appeared to be a function of local district wealth. This inequitable situation needs to be corrected." Perhaps the greatest task for educators in today's classrooms becomes how to effectively teach all children and how to overcome inequalities for children of poverty.

In this task, the school should serve as a counterbalance to poverty by leveling the playing field for all students. For accountability purposes, comparison based upon the poverty status can reveal the presence of such counterbalances or some other value added measures. After the study of effective schools, Ron Edmonds (1979, p. 15) stated, "Inequity in American education derives first and foremost from our failure to educate the children of the poor."

Accountability as Reform

Accountability is not new to education. As explained by Cuban (1990), education like governments go through phases of a cycle. In nineteenth century England, the concept that schools should be held more accountability for student learning was implemented as payment-by-results, and schools received funding based upon student testing results (Cauble, 1992; Renaldi, 1992). In the payment-by-results system, school inspectors assessed the performance of students using standardized testing. Schools then received funding based on the level of student performance on the standard examinations. According to Renaldi, the assessment was restricted to reading, writing, and mathematics. Other content areas, such as fine arts and social studies, were dropped from the assessment because of difficulties in measurement. Although educators objected to the payment-by-results systems, the system continued with political support until the end of the century (Cauble, 1992; Renaldi, 1992). Payment-by-results spread to the English colonies in America where oral examinations replaced the written assessments (Cauble, 1992).

By the early twentieth century, accountability in the United States supported the "sound and cheap" efficiency method of education very similar to the English payment-by-results system (Renaldi, 1992, p. 37). The efficiency movement had students grouped by chronological age and maintained a narrow curriculum emphasizing basic skills instruction. With the focus on efficiency, the public accepted high dropout rates (Renaldi, 1992). As a management philosophy, strong financial bookkeeping and basic skills assessment served as an efficient method to ensure accountability of public schools. With

the educational measurement movement, the belief was that whatever existed, existed in a measurable amount. By the 1930's, the testing continued and expanded to measuring pupil performance in addition to that of the school (Cauble, 1992; Darling-Hammond, 1992). Policymakers continued to seek educational systems that are both efficient and effective. During the 1960's, because of concerns over the increased federal funding of education and perceived failure of public schools, the call for accountability reached the national level (Bradley, 1998; Oldham, 1993). In the current educational accountability era, Kaagan and Coley (1989, p.12) state that indicator systems "offer the unique opportunity for state policymakers to affect local education practice in a most efficient way … and to assess the school's direction, mission, and strategy." Although accountability has changed over the past decades, the basic aspects of student testing, narrowing of the curriculum, payment-by-results, and desire for efficiency remain.

Since 1984, public schools have experienced three waves of reform. With the intense reform efforts, states took an increasingly active role in prescribing educational change. From 1983 to 1987, every state in the nation adopted some form of educational reform (Feir, 1995). Burstein, Oakes, and Guiton (1992), Castello (1993), and Bryk and Hermanson (1993) reported that the national passion for educational reform could be traced to the national report, "A Nation at Risk" (1983). "A Nation at Risk" (1983, p. 5) condemned American schools with such statement as, "if an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war."

The first wave of reform following "A Nation at Risk," although focused on state and local level reform efforts, became a national response (Feir, 1995; Johnson, 1997). From standardized testing to standardized expectations, the nation as a whole began to focus its attention and energy on the failure of the American school system. To reach the level of excellence cited in the report, a push for standardization of expectations and educational requirements called for more rigorous courses for graduation, increased student testing, and more accountability (Feir, 1995; Lieberman, 1988; & Johnson, 1997)

Although the first wave of reform efforts focused on standardization, higher standards, expanded testing, and teacher shortcomings, the second wave of reform stressed accountability, decentralization, teacher empowerment, school-based management, and choice. By the late 1980's, in a comprehensive effort to restructure schools to more effectively teach all students in total learning environments, educators emphasized capacity building through professional development and support systems (Decker, 1988; Johnson, 1997; & Saban, 1997). The second wave of reform sought to rebuild strong relationships among all school community members by professionalizing teaching and building a more collaborative school culture (Lieberman, 1988; Saban, 1997).

The third wave, true restructuring, emphasized using the resources of business, industry, and the community to improve education (Johnson, 1997; Saban, 1997). The third wave directed educators into reinventing the classroom by developing new roles for parents and by changing the descriptions of schooling, school leadership, and professional development. The restructuring created a push for community involvement,

site based decision making and vouchers in an effort to reshape the public school (Johnson, 1997; Saban, 1997). In support for total restructuring, Congress backed the Obey-Porter comprehensive school reform demonstration grant program. The comprehensive school reform demonstration grants blanketed the nation, funding research-based total school reform with $220,000,000 in restructuring funds for fiscal year 2000 (USDE, 2000).

Accountability

Accountability means providing the public with adequate information to reach a fair judgment. The new standards being developed by state education agencies provide a strong base for accountability because they offer an examination of expectations, curriculum, and assessment. According to Wiggins (1993, p. 289), accountability includes a "public obligation to a set of standards, criteria, performance targets that seem out of reach but are reachable." This obligation includes a commitment to improve instruction. To improve instruction, however, schools need in-depth information on a continuing basis to increase learning. The accountability process can provide valuable data to schools as they advance toward school improvement.

The accountability system must also promote public understanding that schools might be teaching something that is highly important but not easily quantifiable or testable (Anrig, 1992; Banks, 1994; Clifford, 1995; Goode, 2001; Kane, 1993; Oldham, 1993; & SAAC, 1993). States need to measure broad knowledge and skills at key points as efficiently as possible to develop accountability standards (Anrig, 1992; SREB, 1995;

& SAAC, 1993). In developing accountability testing, Wiggins (1993, p. 266) stated, "the key words are credible and useful-- to the teacher, learner, and parents." To develop effective systems, Anrig (1992) admitted that the best accountability systems are costly but warned that the status quo might be more costly.

Although many (Adams, 1995; Beck, 1995, Clifford, 1995; Ricotta, 1987; & Wilburn, 1996) have found that educators viewed accountability negatively, Streshly and Newcomer (1994) contended that accountability is a permanent feature in American education. In its simplest form, accountability means holding people responsible for their actions. People or agencies are accountable if they must answer for their conduct or performance. White (1994) stated that assessing performance is absolutely necessary for learning. Additionally, he defined under what conditions accountability could become a threat to instruction: when accountability is imposed externally, when it is insensitive to the student, the learning, or the discipline, when the necessary resources are not provided, or when the results are misused.

However in an educational setting, accountability requires a more precise definition that includes who is responsible to whom for what (Brown, 1990; Darling-Hammond, 1989). Wiggins (1998, p. 289) acknowledged that as teachers "we are properly responsible for our impact or lack of one, not merely for their good faith efforts." Additionally, Wiggins (1998) suggests that no teacher can be successful without some accountability. When accountability is measured against a product, standards are relatively easy to establish. When establishing standards for the learning process,

however, the decisions are more difficult and much more political (Banks, 1994; Clinchy, 2001; Kane, 1993; & LaBaree, 2000).

The sheer complexity of the organizational structure of schools limited the ability to hold public education accountable. As a result of the growing polarization over the question of accountability, struggles exist between teachers and administrators, between rival factions of teachers, between superintendents and school boards, between groups of community members and schools, and so on. While division among these groups has always existed, during the 1990's the question of accountability brought an unrivaled intensity to the struggles (Theobald, 1995). Because of past experiences, many educators view accountability as divisive, detrimental, and counterproductive (Adams, 1995; Burke, 1999; Bush, 1999; Hart, 1999; & Herrington, 1993). White (1994, p. 3) stated, "Most teachers at all levels work hard at their teaching, and they regard assessment as an intrusion into their lives, for no good purpose." From this viewpoint, politics and perception can impact the intended outcome of improving student performance.

Based on the findings from several studies (Beck, 1995; Bradley, 1998; Clifford, 1995; Howard, 1998; Oldham, 1993; Powell, 1997; & San Miguel, 1996), accountability has changed the behavior and methods of both teachers and administrators. Although teachers perceived accountability measures as negative (Adams, 1995; Hart, 1999; & Wilburn, 1996), still they shifted instructional strategies, methodology, or content to become more successful in meeting the content or performance standards of accountability systems.

While there are many unique accountability structures, San Miguel (1996) categorized accountability into educational indicators and school delivery systems. With school delivery systems, schools are held accountable for program compliance or a prescribed methodology (Brown, 1990; Bryk & Hermanson, 1993; Burstein, 1992; Darling-Hammond, 1992; Odden, 1992; & San Miguel, 1996). The prevalent system operating in public education in the past decade was the educational indicator system. Within the educational indicator system, key components surface: district standards, related assessments, awards and sanctions, and public reporting practices (San Diego City Schools, 1993; San Miguel, 1996). In an accountability system based upon educational indicators, performance reporting is "the precursor to the determination of rewards and sanctions" (Debray, 1999, p. 9). Policymakers view public reporting as a means to improve student performance by holding schools accountable for meeting the established performance standards and by changing the teaching and the learning process (Bradley, 1998; DeBray, 1999; Kelley, 1998; & San Miguel, 1996). According to the Southern Regional Education Board's 1995 report, any public reporting should link accountability information to local improvement efforts.

For the past decade, the increasing use of such data as the basis for school improvement created the "culture training of school reform" (SREB, 1995, p. 13). If schools are to use accountability reporting as a cornerstone for planning and school improvement, the reported information provided must be fair and understandable to the schools. Beyond fair treatment, educational indicators should play a constructive role in school improvement. In this positive role, according to Darling-Hammond (1992), an

indicator “illuminates educational issues without distorting the educational process or harming students and schools by misusing indicator data” (p. 235). Darling-Hammond also argued that educational indicators should be conceptually and technically sound without political manipulation, thus fair and equitable to all schools. Although she linked educational indicator systems to public policies, she cautioned policymakers of the danger of overusing indicator data. She stated that, “mere publication of indicator data can change behavior in ways that invalidate the indicator” (1992, p. 249).

The relationship between use and purpose is also found in Messick’s (1989) unitary concept of validity. Messick stated that test use always occurs in a cultural setting for the intended purpose of making decisions. However, each decision has social consequences that could interfere with the test’s meaningfulness. Therefore, educators and policymakers must consider the intended and unintended consequences. If this is true, then it becomes imperative that all reported information reflect a valid foundation for campus change. With all public reporting, policymakers and educators must make every effort to ensure that the indicator is fair, valid, and equitable for all schools.

Accountability Choices

To improve the public schools, policymakers have two basic choices in designing an accountability system: school delivery systems to regulate a prescribed methodology for educators or educational indicators to establish quality standards for students (Brown, 1990; Bryk and Hermanson, 1993; Burstein, Oakes, and Guiton, 1992; Darling-Hammond, 1992; Odden, 1992; & San Miguel, 1996). Although two extremes exist, vast

arrays of opinions emerge. Brown (1990) cautions that while there are two extremes of accountability, neither alternative in its extreme is acceptable. In the predominant system of the 70's and 80's, the prescribed methodology system, the superstructure of monitoring was developed to maintain compliance to federal and state regulations (Brown, 1990; Newman, 1997). New regulations were simply added to existing ones, resulting in an increasingly complex system of rules and regulations. This high degree of institutionalization through rules and regulations guaranteed the resources of the system were distributed in a prescribed way and that the procedural requirements associated with fairness and good instruction were followed (Brown, 1990; Newmann, 1997).

Yet, these very rules and regulations often hampered the creativity and productivity of educators in improving student performance (Brown, 1990; Darling-Hammond & McLaughlin, 1995; & Wilburn, 1996). Conversely, a relaxation of rules through reliance on professional or political models can increase the danger for inappropriate use of funds, acts of discrimination, and continuation of ineffective programs (Brown, 1990; Feuer, 1995). Additionally, this superstructure method of monitoring did not produce the student performance or political results expected. In the prescribed superstructure method, the accountability rested on compliance with the system by focusing on what the educator did, not on what the student learned. In this method, a school could be both compliant and ineffective.

The educational indicator system consists of established statistics to measure the quality of education, such as student assessment, dropout, and attendance data. The current educational indicators in Texas, such as the Texas Assessment of Academic Skills

(TAAS), graduation rate, and advanced placement course completion, focus on the student and the student's performance. In this system, accountability is measured by how well the students achieve or perform. On the other hand, the school delivery system seeks to ensure that each student has an opportunity for a quality education. Herein, the opportunity to achieve the established high standards is accomplished by requiring each school to provide the required curriculum, employ certified teachers, offer minimum diploma requirements, and incorporate national content standards. The school is held accountable to fulfill the established requirement. The school delivery system places accountability on the school to provide the educational setting while the educational indicator system places a majority of the accountability on the performance of the student.

As policymakers and the public became more concerned with the quality of public education, there was an increased emphasis on outcome-based accountability (Clinchy, 2001; Garn, 1998; Harrington-Lueker, 1998; Stout, 1990; Oldham, 1993). Bryk and Hermanson (1993, p. 466) state, "Much of the current interest in indictors has an explicit political purpose." According to Kaagan and Coley (1989, p. 22), politicians rushed to develop accountability systems because they were frustrated "at the foot-dragging within the education system on assessment and accountability." This rush to respond, according to Kaagan and Coley (1989), led to the implementation of accountability systems prematurely before states had the capacity to manage such complex systems.

This concern with student performance more specifically focused on the achievement gap between student groups (Harrington-Lueker, 1998; Herrington, 1993; &

Oldham, 1993). Because of this increased interest in the achievement gap, policymakers turned to an educational system that reports performance by demographic group (SAAC, 1993). Herrington (1993) states two reasons policymakers support educational indicators with a strong public reporting component: to rally pressure from the parents and other community members and to motivate all stakeholders to focus energies upon closing the achievement gap. Although research is limited on the actual impact on student performance, Herrington (1993, p. 5) stated that the impact of educational indicators is "substantial and reasonable."

Kaagan and Coley (1989) stated that although states once maintained control through monitoring, rules, regulations, and allocations, now control comes through the educational indicator systems. As the shift occurred from the school delivery systems to educational indicators, the political cry became flexibility with accountability (Newmann, 1997; SREB, 1995). Each district could locally tailor instruction and programs to best serve its students with the condition that the state's performance standards would be met. Politicians and educators alike accepted local control with high accountability as the new national policy. Because the state agencies were not prescribing programs or methodology, districts could be creative in designing local programs that were both effective and efficient. It no longer mattered which methodology was used to reach the state content and performance standards, but only that the standards were reached (Newmann, 1997; SREB, 1995).

As found in the Southern Regional Education Board (1995), accountability assigns responsibility for student performance to the local education agency but grants

the necessary flexibility to the local educators. Even with increased local control and flexibility, concerns do exist. Regardless of local flexibility, local educators cannot meet the required standards without the necessary resources. In order to support improved student performance, Heffron (1998) found that school improvement is only possible when adequate resources are available. Schools must have these resources and support to be held accountable. When schools lack the necessary resources, the district or state must react with increased responsiveness to the local needs for the accountability system to work (Heffron, 1998; SREB, 1995).

Several ideas exist on the development of accountability systems. Because Anrig (1992) believes that accountability is more than testing, he suggests that accountability systems test less and pay "more attention to what schools are doing" (p.35). Because the assessment system rewards the best test takers, Dr. Sternberg (Goode, 2001, p. D7) advises, "the gaps in a system that prizes conventional testing above all else are problematic." Dr. Steinberg also suggests that society broaden its definition of achievement and ability to encompass more than those tested (Goode, 2001).

Streshly and Newcomer (1994) suggest that to establish equitable and effective accountability systems, teachers and staff should be included in the decision-making process prior to implementation. According to the SREB study (1995), ownership, cooperation and consensus on the standards are vital when using accountability with school improvement efforts. Henry (1996) and Bernauer (1997) advocate an alternative accountability system, community accountability. While rejecting the external pressure from state models, the community model aligns educational indicators with local goals

and establishes the community as the center of the assessment development. Herrington (1993) supports a community approach by connecting accountability and school improvement. Both accountability and school improvement have the campus as the center, not the district or state. School improvement happens at the campus level first and best.

DeMouldin and Kendall (1993) suggest an accountability network with a well-established hierarchy of responsibility. This coordinated effort to improve student learning may be the most beneficial in implementing change. In this system, student performance is the central focus of each segment in the accountability network. Ball and Goldman (1997) propose that education now needs a "system perspective" stressing performance, efficiency, and effectiveness similar to a marketplace system (p. 228). Wiggins (1993) suggests an accountability system based on reward entrepreneurial that provides freedom and opportunity for teachers. In contrast, Davies and Williams (1997) and Riegel (1992) emphasize that the accountability system should only compare a school to its past performance, never to other schools, to increase the importance and value of the results.

Policy Implications

According to Kaagan and Coley (1989), "Accountability is hot!" Although that comment was made in 1989, national interest has maintained the focus on accountability into 2000. With Al Gore, Joseph Lieberman, and George W. Bush proposing stronger accountability and more flexibility in federal spending, accountability is a bi-partisan

issue (Brownstein, September 2000). The Republican platform completely embraced Bush's strong federal accountability with increased school choice for parents (Toner, 2000). The importance of accountability in both platforms was reported throughout the campaign. While speaking at Morehouse College in Atlanta, Gore stated, "The phrase higher standards and more accountability…has to be integrated into the fabric of everything we do when it comes to education." (Brownstein, April 2000). On the campaign trail, in comparing the Gore education platform to his own, Bush stated that accountability was the real issue (Neal, 2000). Throughout the 2000 presidential campaign, accountability and school violence emerged as the two top education issues (Balz, 2000; Helfand, 2000; & Henry, 2000).

Again at the national level, President Clinton's farewell address to the nation, January 18, 2001, included educational accountability: "Our schools are better — higher standards, greater accountability and larger investments have brought higher test scores, and higher graduation rates." The message of accountability continued at the 2001 National Title 1 Conference in Phoenix, Arizona. Keynote speaker Mary Jean LeTendre, the U. S. Department of Education's director of compensatory education, demanded that the nation "not to just count each child, but to account for each child."

In an era of increasing review of education, America has seen an "explosion of interest in education indicators" (Bryk & Hermanson 1993, p. 451). Policymakers are reluctant to entrust the future of education to educators without a clear system of accountability to ensure effectiveness (Bryk & Hermanson, 1993; Darling-Hammond, 1992; Kaagan & Coley, 1989, 1989; Oldham, 1993; & San Miguel, 1996). Darling-

Hammond (1992, p. 235) states the interest and implementation of indicator systems have been linked to "governmental monitoring functions and policy agendas, thus enhancing their political as well as educational or scientific interest."

Kaagan and Coley (1989) state that indicator systems will continue to grow across the nation as long as the state leadership continues its involvement in improving America's schools. Darling-Hammond (1992, p. 237) cautions, "Changes in society's view (or even political party in power) can and do influence the nature of data, its reporting, and use." A noteworthy example is the change in the reporting of dropout data. In the past, the dropout rate was not a primary indicator because student dropouts were able to seek gainful employment on farms or in factories. Now because of economic reasons, it is a critical indicator of school success. Such economic ramifications directly impacted educational policy decisions which in turn impact accountability systems (Darling-Hammond, 1992). Kane (1993, p. 3) described this impact of politics on standards and accountability as "unavoidable arbitrariness." In regard to political decision-making, Kane (1993, p. 4) wrote, "These decisions could be changed, and often are changed, when made by different persons, at different times, or under different circumstances."

Policymakers will continue to seek the necessary data collection for policy development and reform efforts as they evolve out of "education management into school management" (Kaagan & Coley, 1989, p. 3). Kaagan and Coley' statements are based on a national survey of state implemented accountability systems. Of the 35 states with a student performance based accountability system, the survey revealed that 25 states had

school level reporting systems. Additionally across the nation, 38 states included some form of comparisons in the reporting process. Furthermore, 25 states connected the accountability reporting system to policy development. With half of the states directly connecting accountability reporting to policy development, the increased state level involvement in school management is likely to continue (Kaagan & Coley, 1989). With the increased state involvement, the need for more standardization of expectations and performance is the clear message that policymakers send to educators at the local level (Kaagan & Cooley, 1989). After establishing such standards, policymakers can easily tie the use of educational indicators directly to student performance in an accountability process (Bryk & Hermanson, 1993; Harrington-Lueker, 1998; & Oldham, 1993).

Kaagan and Coley (1989) provide insight into the prime purpose of state educational indicator systems. Although educational indicators assess the direction and the mission of instructional programs, each possibility exists as "a continuum from providing simple information at one end to labeling schools or districts as academically bankrupt at the other end" (Kaagan & Coley, 1989, p. 11). According to Kaagan and Coley (1989), there are two general applications for indicator systems: to reveal effectiveness of the system to the public and to hold systems accountable for the results. Bryk and Hermanson (1993) state that indicators should be used as an effort in "community education and crafted in such ways that encourage a continuing, broad-based public involvement with education issues." For broad-based involvement (Bryk & Hermanson, 1993), this process cannot be limited to releasing the school report cards to the media because such information does not truly deepen the community understanding

and knowledge about the school. The second application of holding school systems accountable for results is far more than providing simple information to the public because of the implied enforcement and consequences of rewards and sanctions within the accountability system.

Bryk and Hermanson (1993, p. 458) stated, "the rhetoric of the indicator movement is replete with exhortations about how these data will increase accountability and drive new policy to improve education." Additionally Bryk and Hermanson (1993, p. 452) propose that, "indicators are efficacious instruments with which to monitor the education system, evaluate programs, diagnose its troubles, guide policy formulation, and hold school personnel accountable for results - an impressive array of tasks." Thus legislators are not only funding more education programs, they are also funding more accountability (Darling-Hammond, 1992; Harrington-Lueker, 1998; & Oldham, 1993). Bryk and Hermanson, (1993, p. 454) call educational indicators the "newest tool for legislators and administrators to construct rational policies … and the logical next step on a long list of efforts to 'hyper-rationalize' public education." Increasingly, accountability and evaluation are based on student performance as the measure of effective programs. Indicator systems utilizing student performance are viewed as providing the kind of quantifiable data necessary for important funding decisions (Darling-Hammond, 1992; McLaughlin, 1991; Oldham, 1993; & Ramirez, 1992). According to Odden (1992, p. 24), policymakers use educational indicators as "policy levers they can pull in order to improve student performance." San Miguel (1996, p.33) called educational indicators "automatic triggers for specific actions." If educational indicators are used to control

schools through rules, procedures, rewards, and sanctions, "educational indicators hold considerable potential for advancing education, but they also offer much opportunity for misuse and abuse" (Bryk & Hermanson, 1993, p. 453).

To limit the misuse and to produce satisfactory results, Ramirez (1992) believes that there must be clear delineation of roles and responsibilities to maintain the focus on improved student performance. In a caution to policymakers in the collection of data, Bryk and Hermanson (1993) caution that there is a natural temptation to apply a causal relationship whenever data is collected. As a basis for any state system, according to Kaagan and Coley (1989, p. 30), state-level indicator systems need to consider three design features, "availability of school level data, availability of information on the quantity and capability of school staff, and careful consideration of the unit of data collection, analysis, and reporting."

According to Ramirez, states must establish the standards for the collection, treatment, and use of data without regard to any political pressures. Darling-Hammond (1992, p. 236) stated her concern about the use of data was based on the tendency to use data as the "sole basis for making policy and administrative decisions." She goes on to state that educational indicators were developed to be evaluative in nature, not purely informative. Avoiding political pressure and manipulation is an awesome task. However, any accountability system's creditability and acceptance relies upon the fair use of the collected data. Accountability can even be perceived as a barrier to program implementation when the accountability system does not consider the complexity of the school or contextual factors (Bechtel, 1997). "In the long run, a failure to find an

appropriate role for indicator information may lead to an era in which educational 'management by measurement' undermines rather than enhances the capacity of the system to support student learning and school improvement" (Darling-Hammond, 1992, p. 236).

Educational Indicators

As a tool for assessing student outcomes, educational indicators become visible attempts to move beyond rules and regulations to focus more directly on learning outcomes. Indicators provide the context for understanding school operation and interpreting the student performance results (Bechtel, 1997; SREB, 1995; & Accountability Manual, 2000). Educational indicators are based on "simple descriptive statistics of central tendency such as means and proportions" (Bryk & Hermanson, 1993, p. 471). As sets of statistics, educational indicators reveal something about the condition or performance of a campus, a district, a state, or a nation (Bryk & Hermanson, 1993; Brown, 1990; & San Miguel, 1996). Darling-Hammond (1992, p. 235) stated that indicators are the yardsticks used in "monitoring of the health and progress of American schools." While these indicators often measure student performance only as a snapshot, they do provide valuable information about the situation in which the performance occurred. Educators use indicators to examine relationships through the input-process-output model (Kaagan & Coley, 1989).

According to Kaagan and Coley (1989), the reporting of scores alone is not enough; an accountability system must include an understanding of why the changes

occurred. To give any significance to the indicator or the reported results, there must be an understanding of the dynamics of the system (Kaagan & Coley, 1989). The two important decisions regarding any educational indicator system are what specific type of statistic will be used and what will be the standard of comparison (Bryk & Hermanson, 1993, San Miguel, 1996). Each indicator has a point of reference or comparison that could reveal potential problems or concerns (Richards, 1988). Because indicators are statistics revealing the condition of the school, the important next step is making that statistic meaningful by a comparison. Kaggan and Coley (1989, p. 7) state, "Looked at as a number on a barren page, a statistic, however important the phenomenon it measures, has modest usefulness. Only when it is measured in comparisons does it become significant or useful."

In the search for meaningful comparison, Bryk and Hermanson (1993) state educators most often use three levels of comparisons. The first is the comparison to oneself, which is known as the developmental model. In this model, a school is compared to itself in terms of growth over years. This model is in agreement with Davies and Williamson (1997) and Riegel (1992) that a school should only be compared to its past performance, never to other schools, to increase the importance and value of the results. The next comparison model compares one unit to another unit, which is known as the racehorse model. The unit model compares state against state, school against school, and student group against student group by providing a relative placement for the different units or groups. Strictly interpreted, the racehorse model is primarily interested in who finishes first, regardless of where they started. However, this model does not allow for

contextual factors such as poverty levels, mobility rates, minority status, or even past performance in the comparison. The third model is the comparison with an external standard, which is known as the educational goals model. In the goals model, schools are compared with established goals or standards that represent high levels of quality. The goals model sets the level of excellence that each school strives to reach, while evaluating the school's level of success. Often accountability systems use a combination of the three models to fully evaluate the campus performance and growth. In Texas, the Academic Excellence Indicator System (AEIS) utilizes some aspects of all three models in a comprehensive combination system. The developmental model includes the reporting of multiple years of student performance and rewarding for continued success. The racehorse model compares each campus against a campus group. Finally, the goals model compares the campus performance to a state established performance standard of excellence. In a comprehensive model like Texas, the use of comparison data is very important because it appears in each level of accountability.

As found in the SREB study (1995, p. 7), educators were greatly concerned with the comparing of educational indicators and with the lack of and quality of "comparable data." Even when reported by student groups, Castellano (1993) found similar concerns by campus and district administrators with the use of comparison data. Administrators had great concerns with the use of the Illinois student performance based accountability system, the New Illinois Recognition System. Both superintendents and campus principals had strong concerns over the use of comparison data for special education and limited English proficient students.

Beck (1995, p. 165) also found concerns over unfair comparisons: "Public perception of our low scores translates into messages like 'This school system is a poor one,' or 'Teachers in this system are lousey [sic]'." When comparing schools, Beck found that teachers related higher scoring schools as more effective schools, which had resulted in a 22% increase in school transfer requests by teachers. However, comparisons can provide more information than the mere snapshot of student performance (Bryk & Hermanson, 1993; Davies & Williamson, 1997; Riegel, 1992; & SREB, 1995). In interpreting comparison data, Wiggins (1993) emphasized that both inputs and outputs must be considered in evaluating schools. White (1994, p. 267) stated, "The output of an institution …does not really tell us much about its educational impact or educational effectiveness in developing talent. Rather outputs must always be evaluated in terms of inputs."

Educational indicators "help define the context in which schools and districts operate" by providing valuable information about the situation in which the performance occurred (SREB, 1995, p. 4). Bryk and Hermanson (1993) suggest the wise use of indicators includes reporting student performance within the context to understand adequately the importance of the data. Bryk and Hermanson (1993, p. 462) proposed a more all-encompassing approach: "For a system to achieve completeness would require an explicit representation of the diverse aims of school, the means-end linkage for each of these aims, and the interrelations … that might exist among these multiple, interwoven educational processes." They also stated that policymakers need to consider student learning, pedagogy, instructional practices, organizational resources, and leadership.

Bryk and Hermanson (1993) visualized the school as a tapestry rich in texture and vibrant color, with some strands clearly exposed to all, while others are hidden except to the expert eye. They cautioned that most educational indicator systems depend only upon the strand visible to all, not the hidden ones. “Great danger posed by indicators is that we grow to value what we measure over what is too invaluable to be measured” (Bryk & Hermanson, 1993, p. 476).

Others voice additional concerns. Because standardized tests are racially biased, the use of tests has harmed African Americans as a group (Jencks & Phillips, 1998). They caution that the use of such assessment continues the negative stereotype faced by all African Americans. Allen (1998) revealed that factors, such as reading inaccuracy, personal opinion, life experiences, question type, and inaccurate comprehension, often limited the true assessment of student performance. Additionally, Allen concluded that a student’s ability could not be measured simply by correct or incorrect responses on a standardized assessment. With such concerns, many have recommended a review of the use and interpretation of standardized assessments and the need for more qualitative measures for assessing student performance (Allen, 1998; Jencks & Phillip, 1998).

Components of Educational Indicator Systems

In providing the context for student performance, educational indicators frequently address all areas of education from the student level to the district level (Brown, 1990). Brown describes the key components of the “accountability loop” (Brown, 1990, p. 4) as the key actors, goals, resources, pre-determined standards, and

rewards/sanctions. The key actors include all stakeholders, from the teacher to the legislators, involved in creating or receiving the responsibility for student learning. The goals clearly delineate what is to be accomplished, such as increased attendance, higher test scores, or raised graduation requirements. The access and control of the necessary resources include materials, supplies, personnel, community resources, and decision-making authority. The predetermined standards establish the objectives, benchmarks and targets necessary to evaluate accomplishments. Rewards and sanctions, such as promotion, salary increases, increased responsibility, loss of control, further training, reprimand, or termination, clearly establish the response to the school's success or failure.

To further describe accountability systems, Brown (1990) used the term "system indicators" as those that include input, process, and output variables. Systems indicators frequently address all areas of education from student performance to district policy. Because these indicators are believed to impact the quality of education and student learning, they are often included in school report cards. The term system indicators include input (curriculum, funding, teacher quality, and facilities), process (decision making, planning, and due process), and output (school report cards and student performance) variables. Input components focus on adequacy and equity of school resources through a review of curriculum guidelines, revenue, teacher certification, and categorical funding for special populations. The process component focuses on inclusion, policies, and program management through a review of the inclusion of staff members in district decision-making, required planning elements in funding regulations, and legal due

process requirements. The outcome component focuses on results of minimum competency testing for students, standardization of achievement tests, campus report cards, and district performance reports. The emphasis continues to shift from input and process measures to student outcome measures. An accountability system based on test scores, limited to only output measures, would totally disregard the contextual input and process variables. Brown (1990) advocates that an effective accountability system should address all three components- input, process, and outcomes- not simply test results.

In contrast to Brown's interaction loop, Newmann, King, and Rigdon (1997) place the school in the center where it is reacting to the external pressures. In this accountability model, the school is clustered around four key features: information on performance (test scores), established standards, consequences (rewards and sanctions), and influential agent (a state department of education). This model provides an organizational structure based on customer satisfaction where there is a service provider and patron. However, in this model, the patron has the power to reward, punish, or if necessary, replace. Newmann, King, and Ridgon (1997, p. 43) state, "The assumption is that teachers will try harder and become more effective in meeting goals for student performance when the goals are clear, when information on the degree of success is available, and when there are real incentives to meet the goals." For this accountability restructuring model to be successful, Newmann, King, and Ridgon also emphasize that the organizational capacity must be supported by effective coordination, shared commitment, and collaboration in support of student learning.

Both models include the key issues of standards, student performance, and a system of rewards and sanctions. In comparing the two organizational systems, by using a cooperative curving cycle, Brown emphasizes the importance of the total involvement of and interaction between key actors in the accountability loop. The Brown model portrays action flowing cyclically from one component to the next while the model by Newmann, King and Rigdon (1997) places the school in the center. The school is reacting to, not interacting with, the external pressures.

The similarities between the models, standards, student performance, and a system of rewards and sanctions, reflect three of the basic components of accountability systems. Oldham (1993) added professional development to these three components to link accountability to campus improvement. Although in most models, reporting is included in the rewards and sanctions component, the Student Achievement Accountability Committee (SAAC, 1993) highlighted reporting as a separate component. The SAAC selected standards, assessments, recognition and interventions, and reporting as the major components of an accountability system. After completing a state-by-state comparison of accountability systems reported in Newsweek, McGinn (1999) enumerated the components as assessments, report cards, ratings, rewards, assistance, and sanctions. McGinn (1999) found only two states, Iowa and Nebraska, reported no accountability measures implemented at the state level. The remaining 48 states reported implementing at least one of the accountability measures. Assessments and school report cards were the most prevalent across the nation.

Public Reporting

When the only thing that matters is student performance, public reporting becomes a very political issue (McGinn, 1999). There are two types of school report cards currently provided to the public, the individual and the compiled report card (Jaeger, Gorney, & Johnson, 1994). The individual school report card is designed to inform parents about the school that their child attends, whereas the compiled school report card includes comparison statistics for several schools. For example, one individual school report card, the AEIS, is released to the public each fall.

The concept of releasing the school report cards to the public is not new. Although the earliest school report cards in the United States were prepared for school administrators around the turn of the century, their renewal began in 1969 with the "Columbus School Profile: A Report of the Columbus Public Schools to the Community" (Jaeger et al., 1994). The first attempts at school report cards included facts and figures about the state and local school district to provide a contextual profile for the community (SREB, 1995). The data profile emphasized only input indicators while providing little or no measure of quality of the instructional process. According to the SREB (1995, p. 4), the first school report cards provided many "pages of little numbers" reflecting little about student performance.

Most districts are required by the state board of education, state legislature, state agency, or local administration to submit information or to publish a performance report card (Gaines, 1992; Jaeger et al., 1994; & Oldham, 1993). The school report cards have moved from emphasizing a school's demographic and contextual factors, which are often

viewed as constraints or limitations, to a model that emphasizes student learning (SREB, 1995). Although the first school report cards focused upon the social and economic context indicators, the movement is now on results with a clear link to school improvement efforts (SREB, 1995).

As found in the Western Michigan University Research Center, the most frequently reported information included student performance assessment, state or achievement, staffing and student demographic characteristics, and course and program offerings (Jaeger et al., 1994). Also in this study, the least frequently reported was school success, school environment, student services, and student engagement factors. Jaeger, Gorney, & Johnson (1994) define school success indicators as graduation rate, promotion rate, and special honors and accomplishments. School environment indicators refer to crime and vandalism statistics, which show valuable information about the learning climate. Student services indicators reveal the availability of health and counseling services and extracurricular opportunities. Student engagement data, relating to the number of suspensions and expulsions, is rarely included in the school report cards. As found in this study, parents wanted information in all categories in order to truly evaluate student success (Jaeger et al., 1994).

According to the SREB report (1995, p. 7), "the best school report cards focus on the progress of local improvement plans; they clearly identify goals and measures of acceptable progress; and they report on student performance in those terms." After analyzing 558 school and 74 district report cards and interviewing 166 parents in two states, Jaeger, Gorney, and Johnson (1994) provide guidelines for developing effective

report cards. Among their recommendations, Jaeger, Gorney, and Johnson (1994) advise providing information on school environment, program offerings, and staffing as a minimum. Also, school report cards should include historical data of three years or longer. Current year information should be displayed in the context of averages with comparison data such as district or state averages.

If designed effectively, Jaeger, Gorney, and Johnson (1994, p. 45) state, school report cards can inform parents about the effectiveness of the community's public school so the public will know "what is right about the local schools and what needs improvement." According to Jaeger, Gorney, and Johnson (1994, p. 42), "school report cards are public statements of the condition of individual schools and the results of their education program." The report cards are designed to provide valuable information on the overall effectiveness of the instructional program and to assist in decision making and planning. Although educational indicators represent valuable information used in accountability systems, educational indicators alone do not constitute the accountability systems (SAAC, 1993).

The Student Achievement Accountability Committee (1993, p. 22) contends that, "a good accountability system is one that selects the right indicators of the performance and reports them to the right audience." In determining the right information and right audience, Davies and Williams (1997) and Jaeger, Gorney, and Johnson (1994) state that parents want to know how well the schools are meeting the community's expectation of student performance. Additionally, public reporting can serve as a marketing strategy to keep education at the top of the local community's agenda by providing an annual status

report. Making known to the public the educational needs and highlighting the inadequacy of current school funding are just two of these strategies. In an era of limited revenue and growing demands in education, increased school expenditures seem more acceptable if officials tie them to the expectation of improved student learning (Jaeger et al., 1994).

Although parents want to know how well their child is performing, does this desire to know extend to the campus or the district level? While this knowledge is perceived to be important, Bradley (1998) and Herrington (1993) reported that parents do not understand or pay attention to school report cards reflecting either campus or district information. Albeit parents are concerned about their child and ask if the school is meeting their family's needs, this concern did not extend to the campus or district level. In both studies, parents did not view the public reporting as a critical issue in evaluating their child's school. Bradley and Herrington based these conclusions on interviews and statements from advocacy groups, educators, community members, and parents. In contrast, the Western Michigan University Research Center conducted a fifteen-month study and found that parents wanted and valued school report cards (Jaeger et al., 1994). By a ratio of two to one, parents preferred the more complex, longer reports to the short, abbreviated report formats.

While parents did not view school report cards as valuable (Bradley, 1998; Herrington, 1993), do teachers and administrators have different thoughts? Bradley (1998) conducted a case study of three high schools to examine the reaction to the first New York School Report Card. Bradley concluded that student performance was not

directly impacted by the report card release, but some of the state's purposes in releasing the school report cards were accomplished. The school report card did inform the public about achievement, provide an impetus for public discussion, and result in a greater sense of accountability for teachers. Perhaps the most significant finding reported by Bradley was the importance of the school report card as a document for improving schools and for directing school reform efforts. The school report card served as a driving force in professional development planning for teachers. Campus administrators believed teachers were forced to look at student performance and to challenge their students. Although teachers reported using the student performance data to analyze and measure achievement growth, teachers perceived the report cards as having little or no impact on teaching and instructional practices. Consequently, teachers did not connect professional development or the analysis of student performance as impacting instruction. While the teachers viewed the report card as important for school improvement efforts, they believed the assessments and improved standards were the major change agents.

Likewise, Howard (1998) found that teachers perceived that the state accountability system did not impact instruction, although many instructional changes were reported. The teachers stated that they changed or modified instruction as a result of increased demands from the campus principal. The teachers viewed the campus principal as the change agent, not the accountability system or themselves. In regard to the state assessment, teachers did not think that the instructional practices had a direct impact on student performance on the standardized assessments.

In Dade County, Florida, Herrington (1993) found that campus level administrators believed that reporting performance data by race and ethnicity had a negative impact on student performance efforts. Administrators believed that reporting by race and ethnicity highlighted the underachievement of minority students. Beck (1995) found that educators felt additional pressure from high-stakes testing and accountability. The greatest pressure appeared to come from the local school boards as they established local expectations for improvement. Through open-ended questioning, Beck found that a majority of the reported negative consequences for low performance came from local school boards. For superintendents, pressure was rank ordered from local board, to community and parents, to politicians, and last to the media. Only for large urban areas was the media a major source of external pressure.

Similarly, Adams (1995) conducted a perception survey of principals and fourth grade teachers in Education Service Center Region V, based in Beaumont, Texas. Adams found that campus-level educators and teachers had an overall negative perception and expressed a lack of confidence in the AEIS reporting. Adams also found that administrators and teachers agreed that the AEIS reporting had resulted in a narrowing of the curriculum to the objectives tested to the exclusion of important content. Additionally, principals and teachers also agreed that the AEIS system had also limited student creativity. Agreeing with campus administrators of Dade (Herrington, 1993), the principals and teachers in Region V saw that public reporting of student performance polarized teachers, principals, and the public. On each issue surveyed, teachers were

more negative than principals. Although various stakeholders viewed school report cards differently, they held their viewpoints strongly (Adams, 1995; Herrington, 1993).

In contrast to the campus-level staff, Herrington (1993) found that district level administrators did not share this negative perception of public reporting. The district and central office administrators were found to perceive the reporting as necessary to assure equal accountability and equal expectations for all students. Even though Bradley (1998) and Beck (1995) reported limited or no impact on student performance, top administrators viewed the performance reporting as informative, important, and useful for internal use in campus improvement efforts.

Likewise, Cauble (1992) found Texas superintendents viewed the AEIS report as a valid measure for school success and beneficial in planning. Successful principals active in campus improvement efforts were data driven and embraced accountability (Cavazos, 1999). Bennett (1997) and Birt (1998) also found an increase in teacher engagement concerning student performance and instructional methodology as a result of campus evaluation and accountability systems.

From the onset, improved student performance was the intended outcome. Because public reporting is one of the most costly and difficult responsibilities of accountability systems, Ramirez (1992) cautioned that the process of gathering and reporting of information could intrude on the schools' primary purpose of instruction and learning. Although the use of school report cards is costly and political, there is a strong connection between accountability, public reporting and school improvement. In the SREB study (1995, p. 9), a legislative staff member stated, "The whole reporting thing

intertwines into everything else. All of a sudden you are asking significant questions about whether you have a valid curriculum, whether it is really working, whether it is what you want children to be able to do. You really can't disengage reporting from everything else."

Summary

Accountability systems impact the educational process in both positive and negative ways. Although the goal of any accountability system is to improve student performance, often the political reality impedes appropriate policymaking. In designing accountability systems, policymakers select from the basic components of standards, assessments, reporting, rewards and sanctions. Within the reporting component, policymakers and educators must use caution to maintain a fair and balanced indicator system. In the reporting of student performance, contextual factors provide a more complete view of school success. Often the use of comparisons adds to the analysis of student performance. When assessing student performance, poverty and other student demographics provide additional information in the interpretation of data. Because poverty has and continues to impact student performance negatively, evaluators need to establish fair comparisons in an effectively accountability system.

CHAPTER III

Methodology and Design

Introduction

This study examined the impact of poverty on comparable improvement rankings for elementary campuses in Texas by reviewing the differences between the current sorting process and an alternative sorting process. Comparable improvement, a campus only measure, "shows how student performance on the TAAS test has changed (or grown) from one year to the next and then compares that growth to that of the 40 schools that are demographically most similar to the target school" (TEA Glossary, 2000, p. 6). This year-to-year growth is computed using the Texas Learning Index (TLI). The TLI, a campus only measurement, is computed using students who can be matched to their prior test score for a comparison.

A unique comparison group of 40 campuses was constructed using six demographic characteristics: the percentages of low income, white, Hispanic, African American, limited English proficient (LEP), and mobile students. This comparable improvement process constructs the comparison group each year, using the comparable improvement characteristics in the order of dominance. To determine the order of dominance for each target campus the six characteristics are ranked from highest to lowest percent. The most dominant characteristic is used to select the 100-campus pool most similar to the target campus. After the initial 100 campuses are selected, the process shifts from selecting campuses to removing campuses from the pool based on the

remaining characteristics. The process continues until 40 of the initial 100 campuses remain in the comparison group.

Instead of using the most dominant characteristic as in the current process, the alternative process used the percent of students in poverty as the initial sorting characteristic. The order of dominance was from highest to lowest for the remaining student demographics. Using poverty as the first sorting characteristic ensured that the initial 100 campuses in the campus comparison group would be the most comparable in regard to the percentage of low-income students enrollment. From the initial 100 campuses, the campus group was refined by the remaining student demographic characteristics in the order of their dominance.

The primary comparison between the current and the alternative process was the potential for change in quartile placement. To further understand the impact of poverty on comparable improvement, the analysis included changes in group average Texas Learning Index changes, award eligibility, and campus grouping. The following four research questions guided the analysis:

1. What are the differences in quartile placement based on the current order of dominance process when compared to an alternative process using the percentage of poverty as the initial selection characteristic?

2. What is the difference in the campus TLI average growth and group TLI average growth based on the current order of dominance process when compared to an alternative process using the percentage of poverty as the initial selection characteristic?

3. What changes occur in award eligibility for comparable improvement additional acknowledgement when the current order of dominance process is compared to an alternative process using the percentage of poverty as the initial selection characteristic?

4. What changes occur in campus comparison group composition when the current order of dominance process is compared to an alternative process using the percentage of poverty as the initial selection characteristic?

Sample of Schools

According to the Texas Education Agency (TEA) Pocket Edition, 20 education service centers and 1,183 school districts actively served students in Texas in 1999-2000. The diverse student composition included 576,083 African American; 1,578,967 Hispanic; 1,721,969 white; and 114,764 classified as other. Adding to the diversity of Texas classrooms, approximately 1,955,012 students qualified as economically disadvantaged.

Throughout the state of Texas, there are 7,395 public single, elementary, middle, and high school campuses (TEA website, 2000). Comparable improvement is calculated only at the campus level; therefore, the unit of study was the campus. Additionally

because of specific testing provisions at the high school level and organizational models at the junior high level, the selection of the sample type was important. At the high school level, the Texas Assessment of Academic Skills (TAAS) is only administered at grade 10, thus involving low numbers of students. In determining Texas Learning Index (TLI) growth for the TAAS test, students must be matched with their prior year test scores to calculate the average growth scores. For the 10^{th} grade students, that process means matching their exit level 10^{th} grade TAAS scores to their 8^{th} grade TAAS scores. For 10^{th} graders, instead of an annual TAAS growth score, the growth score becomes a two-year comparison. Occasionally, the comparison becomes a three-year comparison for students retained in the ninth grade. If students cannot be matched to their prior TAAS scores, then no average TLI growth can be computed. Because smaller numbers of students were included in the comparable improvement analysis, the TAAS growth score was biased toward less mobile students who are more easily matched.

Also at both the junior high and high school levels, students can be exempted from the TAAS test for high performance on the end of course tests. At the end of Algebra I, English II, Biology I, and United States History, students take an end of course assessment. Prior to the spring 2000 exit testing date, a student could be exempted from TAAS by passing the Algebra I, English II, and either Biology I or United States history end of course tests. Because no TLI average growth score can be computed, these students would not be included in the campus's comparable improvement calculation.

Across the state, local control determines the grade combination or organizational model at the junior high level. Campuses can exist as a traditional junior high school and

serve grades 6, 7, & 8. In more rural settings, campuses are classified as secondary and serve grades 6 through 12. Another option is the middle school concept utilizing a school within a school model where students are served through a family approach with a designated team of teachers. In addition to the multiple campus organizational models, junior high curriculum also provides varied approaches. At the junior high or middle school level, high school courses can be provided for the advanced students, thus impacting student performance on TAAS when compared to other students without this learning opportunity. Therefore for this study, junior high and high school campuses were removed from the sample.

Of the 7,395 public school campuses, the remaining 3,698 elementary campuses received an accreditation rating and comparable improvement ranking. Because the alternative process used the percentage of students in poverty as the primary sorting characteristic, the alternative process would produce no change for campuses that had poverty as the first characteristic under the current TEA process. For this reason, campuses with poverty as the primary sorting characteristic in the current process were removed from the campus sample. The remaining 2,403 elementary campuses were the sample.

Calculation of Comparable Improvement in Texas

Comparable improvement, which compares each campus's growth on TAAS to the 40 schools that are demographically most similar, impacts comparable improvement acknowledgements, Texas Successful Schools Award System awards (TSSAS),

Academic Excellence Indicator System (AEIS) reports, and the school report card (Accountability Manual, 2000). As directed by statute, TEA annually constructs a unique comparison group of 40 campuses using six demographic characteristics in the order of dominance as determined by the percentage of economically disadvantaged, white, Hispanic, African American, and limited English proficient students (Accountability Manual). Although not required by current statute, TEA added the percentage of mobile students to the selection process for the campus cohort. In the current process, the order of dominance is the key component to constructing the 40-campus comparison grouping for comparable improvement.

In construction of the comparable improvement comparison group, special education students are included in the calculations. All calculations are rounded to one decimal place. The student demographics used in the calculation of comparable improvement are taken from the Public Education Information Management System (PEIMS) as submitted to TEA by each district. For all characteristics, except mobility, the district enters student information in October and then submits the data to TEA by mid-December. The third PEIMS submission in June of each year includes the mobility information. All campuses are grouped by campus type as elementary, middle, high, or multi-level. Hence, to select the original 100 campuses for comparable improvement for an elementary campus, only elementary campuses would be potential selections.

After the six characteristics for the target campus are ranked from the highest percentage to the lowest percentage, the most dominant characteristic of the six is used to select the initial 100 campuses that are the most similar to the target campus. After

constructing the 100-campus group, the process continues by only removing campuses that are the most dissimilar, the most distant in either direction, from the target campus. The removal process continues until only 50 campuses remain. As the last step, the 50-campus group is reduced to 40 by removing the 10 campuses with the most dissimilar of the least predominant characteristics of the accountability student group characteristics. The accountability student groups include African American, Hispanic, white, and economically disadvantaged. The final step is taken to make the campus group more comparable by the accountability standards. Because only student groups that are included in the accountability system are used, the percentage of limited English proficient and mobility are not included in the final reduction from 50 to 40 campuses (Accountability Manual 2000).

With the current TEA process, the use of the most dominant characteristic to select the initial 100-campus group is the most important characteristic because no campuses are ever added to the comparison group. Because the 100-campus group is only reduced, the campus group might not be comparable in the remaining characteristics. The following example taken from the TEA Accountability Manual (2000, p. 47) demonstrates the current process for selecting the comparable improvement comparison group:

Elementary Campus X demographics in the order of dominance:
50.3 % African American, 40.4% Economically Disadvantaged, 29.9% White, 19.8% Hispanic, 12.0 % LEP, 15.2% Mobile

Step 1: 100 elementary campuses having percentages closest to 50.3 % African American students are identified.

Step 2: 10 schools from the initial group of 100 are eliminated on the basis of being most distant from the value of 40.4 % Economically Disadvantaged.

Step 3: 10 of the remaining 90 schools which are most distant, in either direction, from 29.9% White students are eliminated.

Step 4: 10 of the remaining 80 schools, which are most distant, in either direction, from 19.8% Hispanic students, are eliminated.

Step 5: 10 of the remaining 70 schools, which are most distant, in either direction, from 15.2% Mobile students, are eliminated.

Step 6: 10 of the remaining 60 schools, which are most distant, in either direction, from 12.0% LEP students, are eliminated.

Step 7: 10 of the remaining 50 schools which are most distant, in either direction, from 29.9% White and / or 19.8% Hispanic students, the less predominant characteristics, are eliminated.

Upon establishing the comparison group, TEA continues the comparable improvement analysis by calculating the campus TLI average growth on TAAS over a multi-year period in reading and mathematics. The TLI average growth is determined by matching each student's TLI score on the current TAAS test to the TLI score on the prior test. Students entering Texas schools from other states, home schools, non-testing private schools, or merely absent on the TAAS test day will not have a prior TAAS test for comparison purposes. Although the TAAS test may have been taken, students who have miscoded social security numbers or multiple personal identification numbers cannot be matched to the prior test. If the student cannot be matched to the prior TAAS test, then the student is not included in the campus calculation of TLI average growth. The TLI growth score for each matched student is then added and divided by the number of matched students to produce the TLI average growth score for the campus.

The 40 campuses are then ranked according to their average TLI growth. Campuses are assigned a separate quartile placement for reading and math. Writing is not included in comparable improvement due to the lack of TLI scores on the state writing assessment. In the reporting of quartiles, a campus scoring in Quartile 4 represents the lowest growth in performance; likewise, a campus in Quartile 1 reflects the highest growth. Specifically, the ten campuses with the highest average growth, in the top 25%, in TLI on TAAS assessment are assigned Quartile 1 placement. The next ten campuses are placed into Quartile 2. The next ten campuses are ranked Quartile 3. As the ranking continues, the ten campuses in Quartile 4, the lowest 25%, reflect the lowest average growth in TLI of the campus group.

TEA has included a series of safeguards in the process of calculating the average TLI growth score. At the student level, in order to increase reliability, certain exclusions are applied. Because the TAAS test is a criterion-referenced test, students either pass or fail. According to TEA (Accountability Manual, 2000), the TLI is not sensitive to extreme high or low scores. When overall performance is exceptionally high or low, growth is not a very reliable or dependable indicator of either performance problems or improvement. Therefore, both high and low performers are excluded from the average TLI growth score (Accountability Manual, 2000). A high performing student is one who scored at 85 or above TLI on TAAS in the prior year. Because the student has already scored at a high level, computing growth is difficult due to the ceiling effects and may result in a negative growth score. Likewise, safeguards to address the unique measurement problems or "floor effect" of the low performing students are needed

(Accountability Manual, 2000, p. 53). A low performing student is one who scored at the minimum TLI score in either in the current or prior year. When comparing the student's score to the prior test score, the resulting growth score is often misleading and does not represent true academic growth.

TEA also established safeguards for the campus. A campus must have a minimum of ten matched students in reading or math to calculate TLI average growth score for either subject area. Also, there must be 24 or more campuses that meet the minimum ten matched students before campuses are assigned a quartile placement. Each quartile usually consists of ten campuses unless a tie in TLI average growth scores exists between two campuses at the quartile separation points or the campus does not meet the minimum number criteria. In case of a tie, a quartile could contain more than or less than 10 campuses (Accountability Manual, 2000).

The Alternative Poverty-First Process for Comparable Improvement

In contrast to the current process of using the most predominant characteristic as shown in the example, the alternative process used the percentage of poverty, as designated by percentage of free and reduced lunch students, as the first and primary characteristic to select the initial 100 campuses for comparable improvement. The alternative process continued to refine the 100 campuses using the order of dominance of ethnicity, limited English proficiency, and mobility to reduce to the 40-campus group. The principal difference between the two processes was the forced use of poverty in the

initial 100-campus selection. Due to the forced choice of poverty as the first and primary characteristic, the 100-campus group would be comparable to the target campus in regard to the percentage of free and reduced lunch students. Because the 100-campus group is only reduced, the campus group might not be comparable in the remaining characteristics of percentage of ethnicity, limited English proficiency, and mobility.

Using the same campus example, the following will demonstrate the changes in the alternative process:

Elementary Campus X demographics in the order of dominance:
50.3 % African American, 40.4% Economically Disadvantaged, 29.9% White, 19.8% Hispanic, 12.0 % LEP, 15.2% Mobile

Step 1 (changed to selecting campuses on percentage of economically disadvantaged): 100 elementary campuses having percentages closest to 40.4 % Economically Disadvantaged are identified.

Step 2 (no change): 10 schools from the initial group of 100 are eliminated on the basis of being most distant, in either direction, from the value of 50.3% African American.

Step 3 (no change): 10 of the remaining 90 schools which are most distant, in either direction, from 29.9% White students are eliminated.

Step 4 (no change): 10 of the remaining 80 schools, which are most distant, in either direction, from 19.8% Hispanic students, are eliminated.

Step 5 (no change): 10 of the remaining 70 schools, which are most distant, in either direction, from 15.2% Mobile students, are eliminated.

Step 6 (no change): 10 of the remaining 60 schools, which are most distant, in either direction, from 12.0% LEP students, are eliminated.

Step 7 (no change): 10 of the remaining 50 schools which are most distant from 29.9% White and / or 19.8% Hispanic students are eliminated.

Although the percentage of economically disadvantaged students is second in the order of dominance, the forced choice of poverty as the first and primary characteristic

used the 40.4% of economically disadvantaged to select the 100-campus group in the alternative process. After the order of dominance was used to complete the comparable improvement process, the alternative process ranked campuses into quartiles based on the campus TLI average growth in reading and in math.

Procedures

Following the established TEA protocol, a data request was submitted to the Communications Division at TEA. Upon approval of the data request, the current and alternative comparable improvement processes were completed. From the current comparable improvement process, the data fields included the campus number, most dominant characteristic, quartile placement for reading and math, percentage of poverty, campus TLI average growth, and campus group TLI average. For the alternative comparable improvement process, the data included the quartile placement for reading and math, campus group TLI growth, and number of campuses on both the current and alternative reports. The quartile ranking as determined by the current process was compared to the quartile ranking from the alternative process for each campus. TEA determined the shift scores using the following formulas: for reading, $QR_{current} - QR_{alternative}$ = reading shift score: for math, $QM_{current} - QM_{alternative}$ = math shift score. The ASCII data file included the shift score for reading and math as well as the other 17 data fields. Regarding the duplication of campuses on both the current and the alternative process, TEA compared the campus groups to generate the number of campuses appearing on both reports. TEA constructed a data file containing the 19 data elements to

analyze quartile placement, campus/group TLI average growth, comparable improvement acknowledgements, and campus grouping for each elementary campus. After completing the current and the alternative comparable improvement process, TEA provided the data file, the text file, and the paper report to the researcher. Using SPSS Base Version 10 software, the analysis was completed.

Analysis

The purpose of this study was to examine the impact of poverty upon the comparable improvement comparison grouping - not to make generalizations to a broader population. Descriptive statistics (Hinkle, Wiersma & Jurs, 1998) describe numerical data through classification and summarization. The primary statistical analysis included frequency distribution, central tendency data (such as mean, median, mode), correlations, and variability.

In response to research question one; the shift score created by TEA reflected the change in quartile placement between the two processes for both reading and math. A shift score of zero reflected no change in quartile placement when using the alternative process, meaning that the campus was not impacted by the change in the sorting process. In calculating the shift scores, the possibility of negative numbers did exist in determining changes in quartile placement. A positive shift score reflected that the campus was ranked in a higher quartile in the alternative process, hence receiving a benefit from the poverty first sorting process. A negative shift score reflected that the alternative process resulted in a lower quartile placement; therefore, using the poverty-

first sorting process penalized the campus. The magnitude of the shift score reflected the number of quartiles improved or lost.

The first analysis was a correlation using Pearson r to determine the relationship between the campuses shift score and the campus poverty level. A correlational analysis was conducted for both reading and math. In addition to reviewing the correlation of the shift score to poverty levels, it was important to note which campuses moved into higher quartiles as well as which campuses were ranked lower. Although seven possible shift scores existed: 3, 2, 1, 0, -1, -2, and -3, the analysis produced five shift scores ranging from -2 to 2. An analysis of the poverty level for campuses experiencing a shift change was also conducted. For reading and math, a frequency distribution was completed for each shift score by percentage of poverty to analyze the impact of the alternative process.

The aggregate shift score for all campuses was reviewed to evaluate the overall impact of the alternative sorting process. If the current sorting process and the alternative poverty-first process were comparable, no change in quartile placement would be expected. Results deviating from zero would denote that the two comparable improvement processes are, indeed, not comparable. Thus, the poverty-first process used in the alternative comparable improvement resulted in changes in quartile ranking.

To answer question two on the impact of poverty on the group average TLI growth, an analysis of the TLI average growth score of each campus was compared to the group TLI average growth score in the current and the alternative process for both reading and math. If poverty did impact the comparable improvement grouping, the two group means would reflect the difference. The difference between the campus TLI

average growth and the campus group TLI average growth was aggregated for all 2,403 elementary campuses. The same process was repeated for the alternative process. The mean, range, and standard deviation were used to compare the current group average TLI growth to the alternative group TLI average. A t test was used to determine if the difference between the two group means was more than could be expected by chance.

For question three, the impact of sorting by poverty on award eligibility was analyzed to see if more high-poverty campuses would be eligible under the alternative process than the current. The comparable improvement additional acknowledgement for high performance is reported on the campus AEIS report, but does not include a monetary reward. The poverty levels of the campuses placed in Quartile 1 with 50 percent or more high performing students under the current system were compared to the campuses meeting both conditions under the alternative system. Scatterplots reflecting the poverty level of the campuses penalized and benefiting for the comparable improvement additional acknowledgement were completed separately for reading and math.

For question four concerning the campus group comparisons, TEA provided the number of campuses reported on both the two comparable improvement reports. For each campus in the sample, the number of campuses appearing on both reports was analyzed using descriptive statistics, mean, standard deviation, range and frequency. The number of campuses duplicated on both reports and the poverty level of the campuses were analyzed using a scatterplot. Because of the curvilinear scatterplot regarding the number

of duplicated campuses and percentage of poverty, a correlational analysis was conducted.

Summary

Chapter III detailed methods and procedures utilized to answer the four research questions. Using data provided by TEA for the spring 2000 accountability reports, this study utilized a descriptive research methodology to analyze the effects of poverty on comparable improvement for elementary campuses in Texas. The data were analyzed by the use of correlations, frequency distributions, and central tendency statistics. A shift score determine the change in quartile placement between the two comparable improvement processes. The specific impact of poverty upon quartile placement, group TLI average growth, comparable improvement additional acknowledgement eligibility, and campus comparison grouping will be described in Chapter IV.

CHAPTER IV

Results

The purpose of this study was to determine the impact of poverty on the comparable improvement ranking for elementary schools in Texas. To understand the impact of poverty, I reviewed the change in quartile ranking, group TLI average growth scores, additional acknowledgement award eligibility, and comparison group composition. With the assistance of the Communication Division of the Texas Education Agency (TEA), I compared an alternative process for constructing comparable improvement comparisons using poverty as the initial selection characteristic to the current process using the order of dominance.

Sample

Of the 7,395 public school campuses in the 1999-2000 school year, 3,698 elementary campuses received an accountability rating and comparable improvement ranking. The reported dominant characteristics for the elementary campuses were 1,574 white, 1,295 economically disadvantaged, 711 Hispanic, 113 African American, 2 limited English proficient, and 3 mobility. Because the alternative process would produce no change for campuses that had poverty as the dominant characteristic under the current TEA process, the 1,295 campuses were not included in the sample. These campuses had poverty levels ranging from 53.2% to 100%. The campus profiles fit two possible scenarios. First, a campus had poverty as its dominant characteristic with either one

predominant ethnicity or multiple ethnicities. Second, the campus had a middle range of poverty with a balance of two or more ethnicities. Although in this scenario the campus had a middle range of poverty, the percentage of poverty was higher than any one of the ethnic populations.

The remaining 2,403 campuses I included in the sample had poverty levels ranging from 2.2% to 99.8%. Of the 2,403 campuses, 1,931 campuses, or 80.34%, had poverty levels higher than 50.2%. In addition, 574 campuses, or 23.80%, had poverty levels at or above 75% and yet poverty was not the dominant characteristic. Of the 2,403 elementary campuses in the sample, 126 campuses did not receive a quartile ranking in reading, and 33 campuses did not receive a quartile ranking in math. If the campus did not have 10 matched students to compute the TLI average growth or had less than 24 campuses in the comparison group, the campus did not receive a quartile ranking. It was possible for a campus to meet the minimum size requirement in one subject but not in the other subject.

Research Question 1

What are the differences in quartile placement based on the current order of dominance process when compared to an alternative process using the percentage of poverty as the initial selection characteristic?

The shift score was computed subtracting the alternative process quartile ranking from current process quartile ranking: for reading, $QR_{current} - QR_{alternative}$ = reading shift score: for math, $QM_{current} - QM_{alternative}$ = math shift score. Using the Product-Moment

Correlation Coefficient (Pearson r), I analyzed the relationship between the reading and math shift scores with the campus poverty level. The Pearson product-moment correlation coefficient (r) is an index that describes the degree to which two sets of data are related to each other and measures the relationship between two variables (Hinkle, Wiersma, & Jurs, 1998). Table 4.1 lists the correlation coefficients between the reading and math shift scores and the poverty levels for the sample. Although both reading and math shift scores were statistically significant at the 0.01 level, no practical significance (Henson & Smith, 2000; Kirk, 1996) existed in either subject. The statistically significant results are primarily due to the large sample size. The correlation was slightly higher for reading than math; however, neither correlation was practically significant. The square of the correlation coefficient, the coefficient of determination (r^2), indicates the amount of variance of one variable associated with the other variable. The coefficient of determination for reading ($.104^2$) was .0108 and for math ($.093^2$) was .008, both of which reflect near zero relationship.

Table 4.1

Correlation of Reading and Math Shift Score to Poverty Level

		% Poverty	Reading SS	Math SS
% Poverty	p. (2-tailed)	1.000		
	N	2403		
Reading SS	p. (2-tailed)	.104**	1.000	
	N	2277	2277	
Math SS	p. (2-tailed)	.093**	.180**	1.000
	N	2370	2277	2370

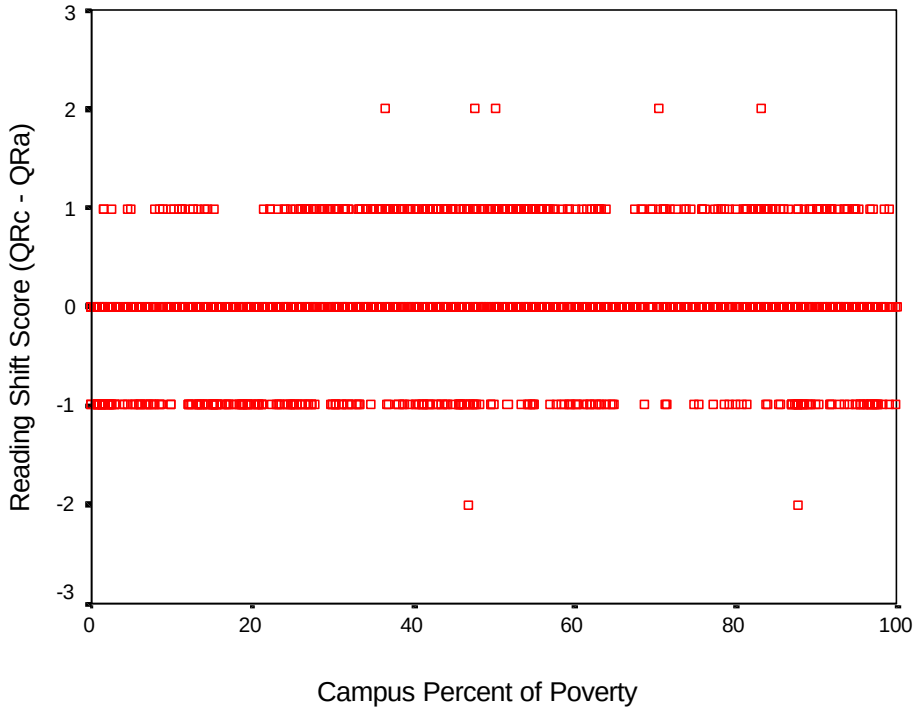
Note **. Correlation is statistically significant at the 0.01 level (2-tailed).

In reading, the distribution of shift scores ranged from -2 to 2. Of the total 2,403 campuses included in the sample, only 2 campuses scored -2 shift score; 290 campuses scored -1; 1,626 campuses scored 0; 321 campuses scored 1; and 5 campuses scored 2. For reading, 126 campuses did not receive a shift score and could not be included in the analysis.

For the 2,277 campuses with a quartile ranking, a scatterplot (Figure 4.1) charted the campus poverty level to the reading shift score. Again, no practically significant relationship existed, as the alternative process affected all poverty levels. Gaps in shift scores existed for various poverty levels. For the -1 shift score, gaps in occurrences were found most prevalent between 60% to 85% poverty. For a shift score of 1, gaps existed below 30% poverty and 60% to 70% poverty.

Figure 4.1

Scatterplot for Reading Shift Score by Poverty Level



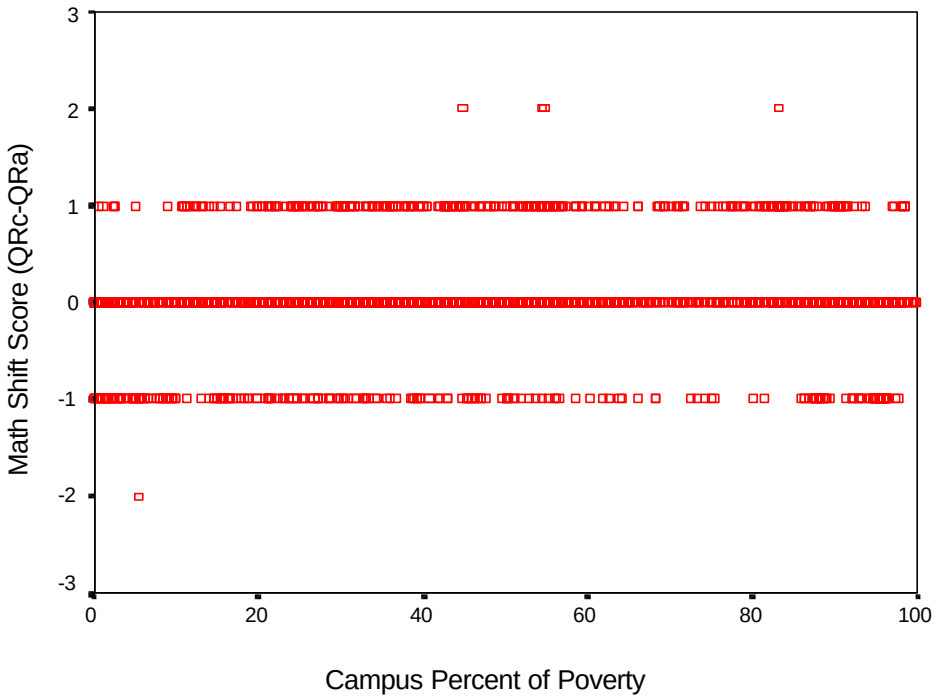
Few campuses received the two extreme shift scores of 2 or -2 reflecting the greatest change. Out of the 2,277 campuses, only seven campuses received either of the two extreme shift scores. The poverty levels and dominant characteristics of the two campuses receiving the -2 shift score, the most penalized campuses, were 46.9% (white) and 87.8% (Hispanic). For the campuses with a shift score of 2, the campuses benefiting the most, the campus poverty levels and dominant characteristics were 36.4% (white), 47.6% (white), 50.2% (white), 70.4% (Hispanic), and 83.1% (Hispanic). The campuses receiving the most extreme shift scores, 2 or -2, represented the poverty levels ranging from 36.4% to 87.8%.

For math, the scatterplot (Figure 4.2) of the campus poverty level with the math shift scores revealed that no practical relationship existed. The regression line crossed all poverty levels for each shift score. Of the 2,403 campuses, 33 campuses did not receive a quartile ranking with the current process and were excluded from the analysis. Out of the remaining 2,370 campuses, one campus scored a shift score of -2, 256 campuses scored -1; 1,760 campuses scored 0; 337 campuses scored 1; and 4 campuses scored 2.

The one campus receiving a -2 shift score had 5.4% poverty and reported white as the dominant characteristic. With a shift score of 2, the four campuses benefiting the most from the alternative process had poverty rates of 44.8% (white), 54.4% (white), 54.6% (white), and 83.3% (Hispanic). Although the one campus penalized was a low poverty campus, the campuses benefiting the most had middle poverty levels.

Figure 4.2

Scatterplot for Math Shift Score by Poverty Level



As a final review, shift scores were summed to evaluate if the alternative process resulted in more campuses ranked in higher quartiles. If the two comparable improvement procedures were comparable, no change in quartile placement would be expected. Aggregate positive or negative scores would reflect an overall change in quartile ranking. After adding all the shift scores for the sample, the reading shift scores summed to 37. More campuses experienced a positive shift than negative shift. The average shift score was 0.02. With such a large sample of campuses, the average shift score revealed a slightly positive change. For math, the shift scores summed to 87, 42% higher than the reading total. The average shift score was 0.04. The alternative process also seemed to benefit the sample and resulted in a more positive shift in math than in

reading, although the average shift score demonstrated only a slightly positive benefit for reading or math.

Research Question 2

What is the difference in the campus TLI average growth and group TLI average growth based on the current order of dominance process when compared to an alternative process using the percentage of poverty as the initial selection characteristic?

To answer research question 2, I compared the mean, range, and standard deviation of the difference of the TLI average growth of the current comparable improvement process to that of the alternative process (Table 4.2). If the alternative process produced a more comparable grouping, the standard deviations may be more compact, reflecting less variability.

Unlike quartile ranking, the TLI average growth scores for the group are computed for each campus group regardless of whether the minimum number requirement was met. Therefore, I included all 2,403 campuses in the sample in the analysis. Although a campus could have had a drop in TLI growth, no campus groups had a negative TLI growth score. The range of TLI average growth scores was 3.96 for the current process and 4.45 for the alternative for reading. For math, the range of TLI average growth scores for math was 2.85 for the current process and 3.37 for the alternative. More variance exists between the campus TLI average growth score and the group TLI average growth score in the alternative comparable improvement process than in the current process. Instead of a more compact grouping, the alternative process

produced more fluctuation. This increased variance in the alternative process was found for reading and math. With the alternative process, the group TLI average growth score ranged from 3.35 to 7.80 for reading, and 3.84 to 7.21 for math as compared to 3.64 to 7.60 for reading and 4.06 to 6.91 for math in the current process.

Table 4.2

Current and Alternative Group Averages by Subject

	N	Minimum	Maximum	Mean	Std Deviation
Reading: Group Avg. TLI Growth Current	2403	3.64	7.60	5.47	.58
Reading: Group Avg. TLI Growth Alternative	2403	3.35	7.80	5.45	.71
Math: Group Avg. Current	2403	4.06	6.91	5.22	.44
Math: Group Avg. Alternative	2403	3.84	7.21	5.20	.49
Valid N	2403				

In comparing the means of reading and math for both processes, the group means were slightly higher for the current process reflecting a lower average growth when campuses are grouped by poverty. The mean for reading in the current process was .02 points higher than the alternative process. Likewise in math, the mean for the current process was .02 points higher than the mean for the alternative process group. However, the standard deviation for math was larger for the alternative than the current process. The greatest difference in standard deviation was in reading with .71 for the alternative and .58 for the current process. For math the difference was much smaller with .49 for the alternative and .44 for the current. Although the campuses were more comparable as far as poverty, there was a difference between the alternative and current group means.

To examine whether the means of the current and alternative process differed by more than would be expected by chance, a two-tailed, paired sample *t* test with 2402 degrees of freedom was conducted. For reading, the *t* critical value (Hinkle, Wiersma, & Jurs, 1998) was 1.960 for the sample size with .05 level of confidence for a two-tailed test. The computed *t* value for reading was 1.584, which was less than the critical *t* value. For math, the *t* critical value was also 1.960. The calculated *t* value for math was 2.426, which was greater than the critical *t* value. Regardless of the variability for reading, the *t* test did not reveal any statistically significant difference between the two means. For math, this difference between the two means was more than what would be expected by chance reflecting a statistically significant difference between the current and alternative process.

Research Question 3

What changes occur in the comparable improvement additional acknowledgement eligibility when the current order of dominance process is compared to an alternative process using the percentage of poverty as the initial selection characteristic?

To receive the comparable improvement acknowledgement for high performance, a campus must be ranked in Quartile 1 and have 50% or more of the students scoring above 84 TLI on TAAS for that subject. This recognition is designated on the campus AEIS report. No financial rewards are attached to this comparable improvement award. A campus can receive this award for reading only, for math only, or for both subjects. If a

campus is not ranked in a quartile for a subject, the campus cannot be considered for the comparable improvement award for that subject.

Campuses Penalized by the Alternative Poverty First Process

Reading.

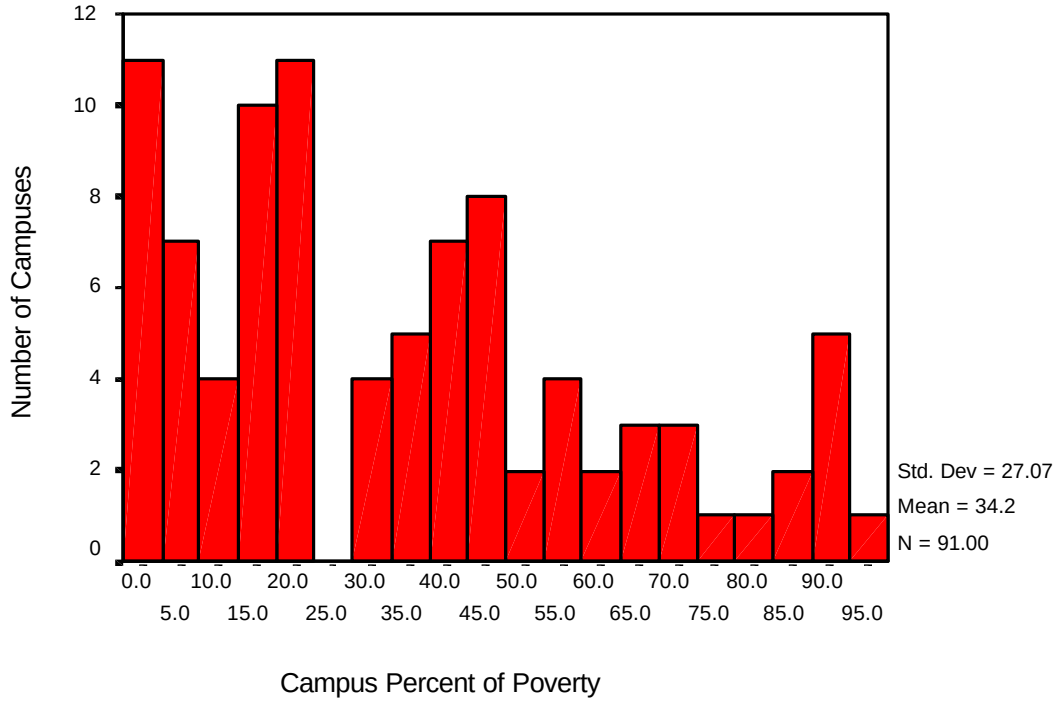
For this analysis, I reviewed the campuses eligible in the current process but ineligible in the alternative process, thus resulting in the campus being penalized. To be penalized, a campus had to have a negative shift score as a result of the alternative process. Of the campuses, 91 campuses met the following three conditions: ranked Quartile 1 in the current process, ranked lower than Quartile 1 in the alternative process, and reported 50% or more of the students scoring above 84 TLI on the last reading TAAS assessment.

The mean poverty level of the penalized campuses was 34.16%. The poverty levels of the penalized campuses ranged from 0.4% to 96.40% poverty. With a range of 96.0 percentage points and a standard deviation of 27.07 percentage points, campuses from all poverty levels were penalized.

In Figure 4.3, a slight tendency to penalize low poverty campuses was found. Campuses with 50% poverty or higher were less likely to be penalized by the poverty first alternative process. Campuses broadly clustered at two poverty levels, 0% to 20% and 30% to 50%. With 2,277 campuses reviewed, 91 campuses, or 4%, reflected that the alternative process penalized low numbers of campuses.

Figure 4.3

Campuses by Poverty Level Penalized in Reading



Mathematics.

The comparable improvement acknowledgement for math is awarded to campuses ranked in Quartile 1 with over 50% of the students scoring above 84 TLI on TAAS math. The included campuses met the following three conditions: ranked Quartile 1 in the current process, ranked lower than Quartile 1 in the alternative process, and reported 50% or more of the students scoring above 84 TLI on the last math TAAS assessment. For a campus to be penalized in math by the alternative process, the campus would have a negative shift score.

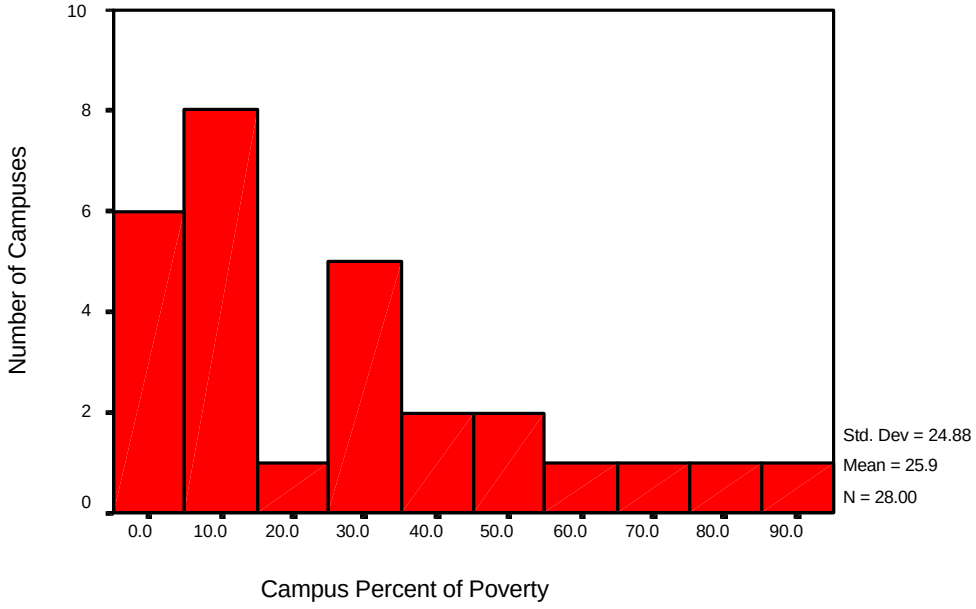
Only 28 campuses were penalized in math (Figure 4.4). The mean poverty of the penalized campuses was 25.91 % poverty. The standard deviation of 24.88 continued to

reflect a broad variance. The range of campus poverty level penalized was 0.4% to 87.20%. With only 28 campuses experiencing negative shift scores in math as compared to the 91 campuses in reading, the alternative process did not produce as much fluctuation in quartile ranking in math as in reading.

The alternative process appeared to penalize low poverty campuses more than high poverty campuses. Campuses with poverty levels of 20% and lower were the most penalized. Above 60% poverty, only four campuses were penalized. With only 28 of 2,370 campuses, the number penalized represents approximately 1.18% of the sample. Although penalizing more low poverty campuses, the actual impact was minimal in regard to overall influence. Of course, the impact was more direct for the 28 campuses.

Figure 4.4

Campuses by Poverty Level Penalized in Math



Campuses Benefiting from the Alternative Poverty First Process

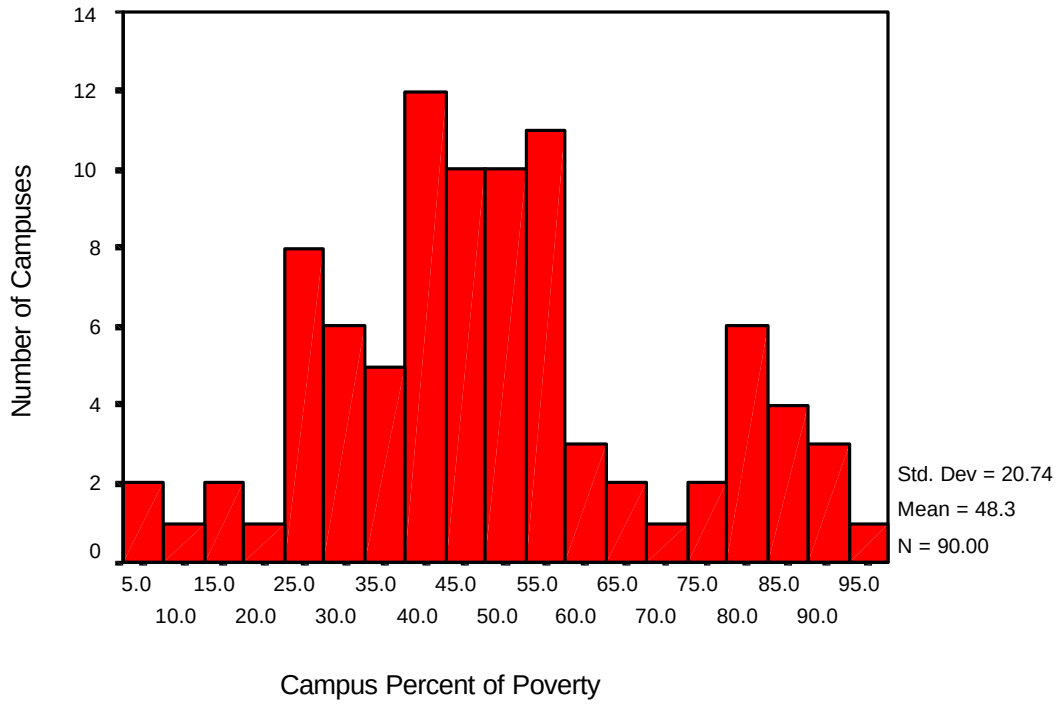
Reading.

Next in the analysis, I reviewed which campuses benefited in reading from the poverty-first alternative process. Campuses had to meet three conditions: ranked lower than Quartile 1 in the current process, ranked Quartile 1 in the alternative process, and reported 50% or more of the students scoring above 84 TLI on the last reading TAAS assessment. Of the campuses, 90 were identified as benefiting, having a positive shift score, from the alternative process. The mean poverty level for the campuses benefiting was 48.32%. The standard deviation was 20.74. The range of poverty levels of the campuses benefiting from the alternative process was 4.50% to 92.70% poverty.

The alternative process seemed to benefit campuses with moderate to high poverty levels more than the current process as demonstrated in Figure 4.5. The figure demonstrates that the campuses with poverty levels ranging from 40% to 60% tended to benefit in higher numbers than either the low or high poverty campuses. With only 90 campuses out of 2,277, the actual number of campuses that benefited from the alternative process was low. Although low in number, the campuses tended to cluster around the middle 40% to 60% poverty level. Two smaller clusters appeared between 25% to 40% and 80% to 95% poverty. Few campuses had poverty levels below 25%.

Figure 4.5

Campuses by Poverty Benefiting in Reading



In summary, comparing the penalized campuses to those benefiting in reading, the mean poverty for the campuses benefiting was 14.17 percentage points higher than the penalized campuses. The alternative process tended to benefit the middle- to high-poverty campuses more than the low-poverty campuses in reading. While the standard deviation was 6.34 percentage points larger for the penalized campuses, the standard deviation of 20.74 still revealed a large variance of poverty level.

Mathematics.

I reviewed which campuses benefited from the alternative process in math. Thirty-two campuses met the three following conditions: ranked lower than Quartile 1 in the current process, ranked Quartile 1 in the alternative process, and reported 50% or

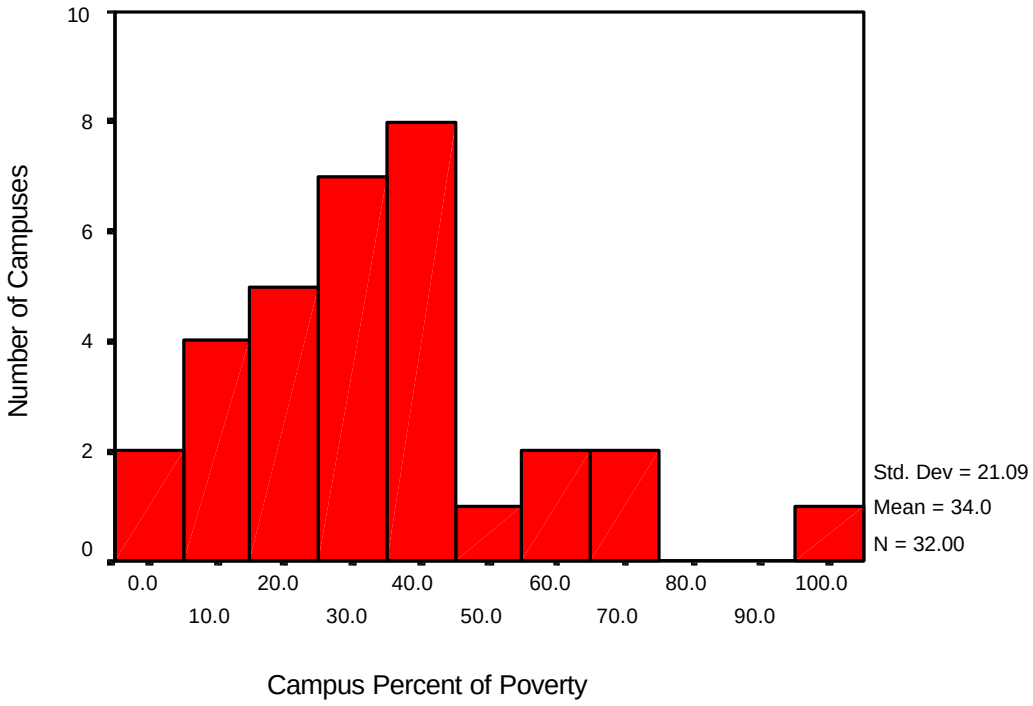
more of the students scoring above 84 TLI on the last math TAAS assessment. Campuses benefiting had a positive shift score as the result of the change in quartile placement in the alternative process.

Thirty-two of the 2,370 campuses met the conditions. The mean poverty level was 33.95%. The standard deviation was 21.09. The range of campus poverty was 97.00 points from 1.10% to 98.10%. Figure 4.6 reports the campuses that benefited from the alternative comparable improvement process by poverty levels. The alternative process seemed to benefit campuses with low to moderate poverty levels, or 10% to 40% poverty.

Although few campuses actually benefited, more low-to-middle poverty campuses benefited from the alternative process. With a standard deviation of 21.09, the variance is still large, crossing most poverty levels. The mean poverty level was 34.0%, which is 8 percentage points higher than the mean poverty of the penalized campuses. Thirty-one of the 32 campuses had poverty levels below 80%. The number of campuses increased forming a patterned stair-step from 0% to 40% poverty. A gap appeared between 80% to 90% poverty in which no campuses benefited from the alternative.

Figure 4.6

Campuses by Poverty Benefiting in Math



When comparing the two groups in math, the mean poverty level for the campuses benefiting from the alternative process was 34%, which is 8.1 percentage points higher poverty than the mean of the penalized campuses. The standard deviation was 21.09, which is 3.79 points less than the standard deviation for the penalized campuses reflecting slightly less variance. In math, the alternative process benefited campuses with poverty levels ranging from 1.10% to 98.10%.

Research Question 4

What changes occur in campus comparison group composition when the current order of dominance process is compared to an alternative process using the percentage of poverty as the initial selection characteristic?

As a result of the data request, TEA generated a data element for the number of campuses reported on both the current and the alternative comparable improvement reports. Although the comparable improvement analysis produces a 40-campus comparison group, the target campus is included in the campus list making a 41-campus report. Therefore, when TEA reported that one campus was on both the current and the alternative reports, the one campus was the target campus. For the campuses with poverty as the dominant characteristic in the current process, all 41 campuses were reported on both reports reflecting the 40-campus comparison group and the target campus. This data element was computed for all 2,403 campuses regardless of either the small number safeguards.

The mean was 4.00 duplicated campuses on both reports (Table 4.3). The standard deviation was 2.87 campuses. The range of number of duplicated campuses reported was 19.0. Only one campus of the 2,403 campuses reported the maximum number of 20 duplicated campuses on both reports. The minimum number of 1 campus on both reports, which reflected the target campus itself, was reported for 346 campuses.

Frequencies for the number of campuses reported on both comparable improvement reports provide additional information. The mode was 3 campuses duplicated on both the current and the alternative reports. Although the maximum score

was 20 campuses reported on both reports, 90.6 % of campuses had 7 or fewer campuses duplicated on both reports and 98.8% of the campuses had 14 or less campuses duplicated on both reports.

Table 4.3

Frequency of the Number of Duplicated Campuses

# of Dupl. Campuses	Frequency	Percent	Valid Percent	Cumulative %
1.0	346	14.4	14.4	14.4
2.0	460	19.1	19.1	33.5
3.0	469	19.5	19.5	53.1
4.0	383	15.9	15.9	69.0
5.0	260	10.8	10.8	79.8
6.0	152	6.3	6.3	86.1
7.0	106	4.4	4.4	90.6
8.0	57	2.4	2.4	92.9
9.0	48	2.0	2.0	94.9
10.0	26	1.1	1.1	96.0
11.0	24	1.0	1.0	97.0
12.0	10	.4	.4	97.4
13.0	19	.8	.8	98.2
14.0	14	.6	.6	98.8
15.0	6	.2	.2	99.0
16.0	8	.3	.3	99.4
17.0	7	.3	.3	99.7
18.0	4	.2	.2	99.8
19.0	3	.1	.1	100.0
20.0	1	.0	.0	100.0
Total	2403	100.0	100.0	

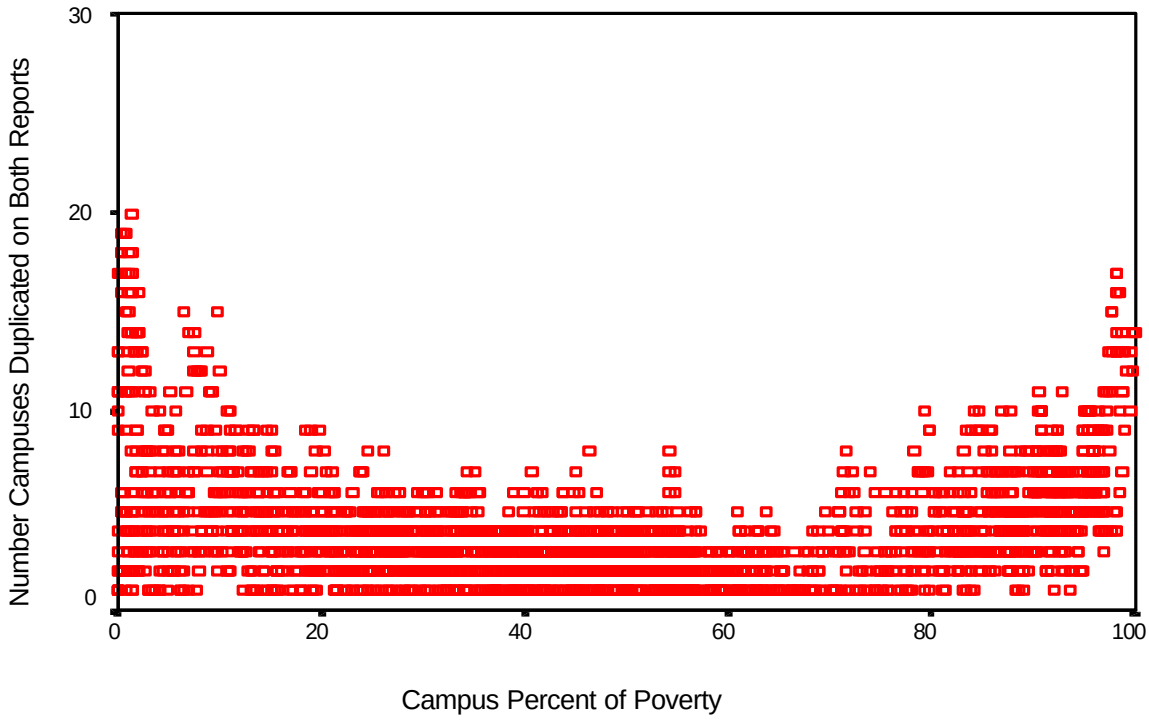
A review of the 346 campuses with one campus, itself, on both reports was conducted. Of the 346 campuses, 91 campuses had a 60% or higher poverty level and 228 campuses had a 40% or higher poverty level. The average poverty level was 46.72%. With the poverty levels ranging from 0% to 93.5%, again impacting all poverty levels.

With a correlation coefficient of .004 (n = 2,403), analysis between the levels of poverty to the number of duplicated campuses revealed no statistically significant correlation.

Using a scatterplot (Figure 4.7), the relationship of the campus poverty level and the number of duplicated campuses on both comparable improvement reports produced a nonlinear relationship. Although nonlinear, a fairly clear curvilinear tendency did exist. I believe that this figure to be the most significant finding reflecting non-comparability. The campuses at either the extreme high or low poverty levels tended to have more duplicated campuses. The campuses in the mid-range of poverty had the fewest campuses duplicated.

Figure 4.7

Relationship Between Number of Duplicated Campuses with Poverty



Because of the curvilinear results, the relationship between the percent of poverty and the number of duplicated campuses was examined more precisely for the bottom and top poverty quartiles. Campuses in the highest poverty quartile, above 71.7 % poverty, and the lowest poverty quartile, below 25.2 % poverty, were analyzed using the Pearson Correlation.

Using the lowest poverty quartile (600 campuses) in the correlation analysis, a low negative correlation of r = -.43 developed between the number of campuses on both reports and the poverty levels of the campuses. With this low negative correlation, campuses with lower poverty levels tended to have more duplicated campuses. Although this is a higher correlation than the total campus group, the coefficient of determination, $r^2 = .19$, suggests that there is a moderate relationship between the two factors.

I completed the same analysis with 602 campuses in the top quartile of poverty (71.7% or above poverty). The result was a moderate positive correlation of r = .52 between the poverty level and the number of campuses on both comparable improvement reports. Campuses with higher poverty rates tended to have more duplicated campuses on the comparable improvement reports. With a correlation of r = .52, the coefficient of determination ($r^2 = .27$) represents a moderate shared variance. The curvilinear relationship between the number of duplicated campuses and campus poverty reflected the greatest impact on the middle poverty level campuses. The middle poverty campuses tended to have fewer duplicated campuses than the campuses with either low or high poverty levels. Figure 4.7 provided the strongest evidence of the incomparability between the current and the alternative process.

Summary

The analysis of the data provided modest evidence of the impact of poverty on the comparable improvement quartile ranking, campus group TLI average growth, and awards eligibility. For each research question, poverty levels were widely variable. Although the impact of the alternative process was spread across all poverty levels, the middle poverty campuses experienced the greatest change. For research question 2, the *t* test comparison of the two group means produced a statistically significant difference between the two processes for math. Additionally for research question 4, the campus grouping, provided a mild to moderate relationship between the poverty-first comparable improvement sorting procedure and poverty.

CHAPTER V

Discussion and Interpretation

The purpose of this quantitative study was to determine the impact of poverty on the comparable improvement ranking for elementary campuses in Texas. The analysis focused on four issues concerning comparable improvement: the quartile placement in reading and math, group average Texas Learning Index (TLI) growth, award eligibility, and comparison group composition. An alternative comparable improvement process using poverty as the initial selection characteristic was compared to the current comparable improvement process using the most dominant demographic. The fundamental issue was to determine whether grouping campuses based on the campus poverty level produced a more equitable and comparable comparison.

The issue of poverty is a mounting concern as the numbers of children in poverty continue to grow (TSSS, 2000). In this study, I sought to expand the knowledge base regarding the impact of poverty on student performance. The findings of this study could contribute new understanding to the perceived influence of poverty on student achievement. An additional outcome of this study was an improved understanding of the state of Texas accountability system and the use of comparable improvement in statewide comparisons. As campuses continue to receive comparable improvement reports and quartile rankings, educators can better interpret and understand the indicator's significance.

Research Questions and Findings

Research Question 1

What are the differences in quartile placement based on the current order of dominance process when compared to an alternative process using the percentage of poverty as the initial selection characteristic?

No practically significant correlation existed between the shift score and the campus poverty levels. The correlation coefficient was slightly higher for reading than in math, however both correlations were negligible. For reading, approximately two thirds of the campuses experienced no change in quartile ranking in the alternative sorting process. For math, approximately three-fourths of the campuses did not change quartile ranking. For reading and math, the middle poverty campuses, 40%-80% poverty, tended to be effected most. The results may also have been impacted by the removal of the 1,293 campuses with poverty as the dominant characteristic in the current process from the study sample.

Using an aggregate sum of shift scores as an overall review of quartile change, the alternative process produced more positive than negative shift scores in both subjects as more campuses moved into higher quartiles. For both reading and math, the aggregate shift scores reflect little overall change when considering the sample size. Although the alternative process produced little change in quartile placement, more positive shift scores were produced in math than reading. One explanation might be that fewer elementary campuses were assigned a quartile ranking in reading under the current

system. Therefore, no shift score could be computed for this study. In the current process, 126 campuses for reading and 33 campuses for math did not receive quartile rankings. This difference of 93 campuses may have accounted for the difference in aggregate shift scores. This lack of quartile ranking could reflect a higher number of student exemptions in reading than in math and thereby reducing the number of matched students. Exemptions would be allowable for limited English proficient and for special education students.

The use of quartile ranking may have minimized the results. It is unknown what change in ranking within the quartile group resulted from the alternative process. With ten campuses (25%) in each quartile, a campus could have moved up or down nine campus positions without changing quartiles. Using the actual ranking within the quartile could have provided a more precise measurement of change. To improve data quality and reduce data entry mistakes, Texas Education Agency constructed the alternative process and the comparisons. Though this action may have reduced potential data entry mistakes, this action also removed any possibility of reviewing change within the quartile.

The results may also have been impacted by the close relationship between ethnicity and poverty (CEAPIR, 1998; Reigel, 1992;TSSS, 2000; Webster, 2001; & Welch, 1995). In the current process, a poor, predominately Hispanic campus would be compared to another predominantly Hispanic campus, regardless of the poverty level. In the alternative process, the poor Hispanic campus would be compared to other poor campuses, which may or may not be predominantly Hispanic. Because of the high correlation between the percent of students of minority status and poverty, the change in

selection process may have produced little variation. White, the ethnicity with the lowest poverty rate, may have had the most benefit from the alternative process. However, because my study focused on the poverty levels of campuses impacted by the alternative process, the impact on different ethnicities cannot be determined.

Research Question 2

What is the difference in the campus TLI average growth and group TLI average growth based on the current order of dominance process when compared to an alternative process using the percentage of poverty as the initial selection characteristic?

Research question 2 was the most difficult question for me to analyze. I determined that the best method to evaluate the impact of the sorting process was to examine the impact on the group means. An analysis of the group mean of the differences between the campus and group averages in both processes provided the necessary comparison. The difference between the campus and group TLI average growth was computed first. Then, the aggregate difference in TLI average growth was calculated for each subject for both processes.

If sorting campuses by poverty had produced a more comparable grouping, then the alternative process should have produced a more compact group average. However, the alternative process produced a broader range of group averages, lower means, and larger standard deviation scores than the current process for both reading and math. In order to better understand the difference in the group means, a two-paired sample *t* test was conducted. For math, the *t* test revealed that the two group means were statistically

significantly different. With three-fourths of the campuses not changing quartiles, the group composition varied enough to impact the group TLI average growth. Although the *t* test revealed no significance in the means for reading, the range and standard deviation for the two groups were notably different.

In this study, I sought to review the impact of the alternative process on the comparison group's TLI average growth. Using the information gathered, I found that the two comparable improvement processes did produce statistically different group TLI average growth scores in math, but not reading. The finding from research question 4 could also have impacted these results. The two comparison groups were practically significantly different; therefore significantly different means would be expected.

Research Question 3

What changes occur in the award eligibility when the current order of dominance process is compared to an alternative process using the percentage of poverty as the initial selection characteristic?

Comparable improvement additional acknowledgement is awarded to a campus for high performance on TAAS. To be considered high performing, a campus must be ranked in Quartile 1 in the subject and have 50% or more of the matched students scoring above 84 TLI. The additional acknowledgment is reported on the campus AEIS report and included in the required public hearing. For this question, the poverty levels of the campuses that either benefited or were penalized by the alternative process were analyzed.

In comparing the number of campuses in each process for reading, 91 campuses were penalized by the alternative process, while 90 campuses benefited. The mean was 34.16% poverty for the penalized campuses, while 48.32% poverty level for campuses benefiting. The alternative process tended to benefit the middle range poverty campuses more than either high or low poverty schools for reading. For math, the alternative process penalized only 28 campuses, while 32 campuses benefited. The mean was 25.9% poverty for the penalized campuses, and 34% poverty level for the campuses benefiting. Although the alternative process impacted more low to middle poverty level campuses, the campuses benefiting from the alternative process had an 8-point higher mean poverty level than those penalized for math.

The alternative sorting process impacted campuses from all poverty levels, but the middle poverty campuses benefited more than high or low poverty level campuses for reading. For math, campuses with 30% to 40% poverty benefited the most. The alternative process mean was 8.1 percentage points higher than the mean for the current process. When compared to math, reading had three times as many campuses impacted by the alterative process. Additionally for reading, the difference of 14 points in the mean poverty levels for campuses benefiting over the campuses penalized demonstrates the impact on the middle poverty campuses. This impact on the middle poverty levels could have been strengthened by the removal of the campuses with poverty as their dominant characteristic. Additionally, the results from research question 4 revealed that the middle poverty campuses had the most change in comparison groups. For both subjects, the total number of campuses benefiting or being penalized was relatively low when compared to

the size of the sample. Because the measure of change was the quartile, the actual change of ranking within the quartile group was not available.

Research Question 4

What changes occur in campus comparison group composition when the current order of dominance process is compared to an alternative process using the percentage of poverty as the initial selection characteristic?

With 2,403 campuses in the sample, the alternative process produced reports with 20 or less duplicated campuses. The number of campuses duplicated on both comparable improvement reports range from 1 to 20 campuses. It was noteworthy that 346 campuses, 14.39% of the total sample, did not have a single campus repeated on both reports, except itself. With a mean of 4.0 duplicated campuses, the average campus had only 3 other campuses on both reports. While 90.6% had 7 or fewer (6 campuses and the target campus) duplicated campuses, 98.8% had 14 (13 besides the target campus) or less campuses reported on both reports.

Although I found no correlation between the campus poverty level and the number of duplicated campuses, a scatterplot (Figure 4.7) revealed that a curvilinear relationship existed. The figure visually displayed the lack of comparability, particularly for the middle poverty campuses, between the two processes. By definition, comparable improvement compares campuses that are the most similar to the target campus. As Figure 4.7 clearly demonstrated, I found the campuses included in the comparison group were not the most and not comparable in regard to poverty.

After assigning quartile placement for campus poverty levels, a low negative (r = -.43) correlation was found for the low quartile poverty campuses. For the low poverty campuses, as the poverty increased the number of duplicated campuses decreased. In the top quartile of campus poverty, a moderate positive correlation (r = .52) existed. For the high poverty campuses, as the poverty increased the number of duplicated campuses increased. The middle range of poverty campuses tended to have the least number of duplicated campuses on both reports. The campuses with the highest and the lowest poverty levels tended to have the highest number of duplicated campuses.

In the current TEA process, the initial 100-campus group must be similar in regard to the most dominant characteristic. However, in regard to the second sorting characteristic, the campuses may be very dissimilar. For the medium poverty campuses, the second sorting characteristic was often poverty. The alternative process forced the initial campus group to be comparable in regard to campus poverty. Research questions 2 and 4 provided evidence that the two processes were not comparable in regard to poverty, however little impact was found on quartile ranking and award eligibility in research questions 1 and 3. Reliability is the consistency of scores between measurements. Although the campus grouping was not comparable in relation to the level of poverty, it proved to be a reliable system of comparison. Few campuses changed quartiles and no campuses moved more than two quartiles. Considering the sample size, the comparison groups provided consistent results with few exemptions.

When reviewed holistically, the alternative process had the greatest impact on the campuses with poverty ranging from 40% to 80%. Because poverty was not the campus's

dominant characteristic in the current, the alternative process tended to provide a more favorable comparison group. Of the 857 campuses with 40% to 80% poverty, the dominant characteristic was limited English proficient for 2 campuses, mobility for 2, African American for 46, Hispanic for 230, and white for 575 campuses. Three possible factors may have influenced the results. First as found in research question 4, the middle poverty campuses had the most change in comparison group. Second, the interaction between the middle poverty campuses and the alternative process could have been fueled by the removal of the 1,295 campuses with poverty as the dominant characteristic in current system. With poverty levels ranging from 53.2% to 100%, the 1,295 campuses could have potentially reduced any impact on high poverty campuses and thereby emphasized the impact to the middle poverty campuses. Third, the high correlation between poverty and ethnicity may have limited the impact of the alternative process.

Discussion

One of the characteristics of poverty is that it knows no color lines (Connerly, 1999). Poverty crosses all ethnicities, all languages, and all ages (Reigel, 1992; TSSS, 2000; & Welch, 1995). As Texas classrooms continue to become more diverse, they also grow poorer (TSSS, 2000). Low-income students tend to perform at lower levels than their more affluent peers. Through the years following the Coleman study (1966), researchers (Adams, 1994; Alexander, 1998; Brown, 1994; Cormier, 1992; Dollinger, 1997; Fortune, 1979; Haetinger, 2000; Hemberger-Coffey, 1991; Hill, 1999; Lamberson, 1989; MacDonald, 1996; Maume, 1998; Perkins, 1992; Pierre, 1994; Plecki, 1991;

Thompson, 1996; Smith, 1999; & Yong, 1987) have found similar results providing a firm foundation for the negative impact of poverty on student performance.

In teaching children to overcome the inequalities of poverty, the school is often the institution to level the playing field for all students. In this task, several studies provide educators encouragement. The Texas Successful Schools Study by TEA (2000), RMC Research Corporation study (2000), Edmonds Effective Schools Study (1979), and the Effective Border Schools Research (1996) found that high minority, high poverty schools can be academically effective. Other studies (Cavazos, 1999; Hill, 1999, McClure, 1999, Powell, 1997; Rowland, 1999; & San Miguel, 1996) found high poverty campuses that experienced high success rates through strong leadership and quality instruction. When schools are effective, all students can learn. These findings continue to support the philosophy of Edmonds (1979) when he stated, "We can, whenever and wherever we choose, successfully teach all children whose schooling is of interest to us. We already know more than we need to know in order to do that. Whether we do it or not must finally depend upon how we feel about the fact, that we haven't done it so far" (School Improvement, 2001).

As policymakers continue to legislate more systems of accountability (Bryk & Hermanson, 1993; Darling-Hammond, 1992; Kaagan & Coley, 1989; & San Miguel, 1996), the focus on student outcomes becomes of utmost importance. The public has a meaningful context for interpreting such student outcome data when states provide a comparison for educational indicators (Bryk & Hermanson, 1993; Davies & Williamson, 1997; Richards, 1988; Riegel, 1992; San Miguel, 1996; & SREB, 1995). Although these

comparisons provide a meaningful context to judge performance, the wrong comparison can be detrimental, divisive, and counterproductive.

To respond to the public's desire for accountability, the state of Texas has developed an elaborate and sophisticated system based on student outcomes. The state system is so detailed that a 200-page technical manual is provided annually. Many state educators view the comparable improvement reporting as complex and complicated, while some view it as unfair and unreliable. Perhaps the greatest outcome of this research will be an improved confidence in the state process of comparable improvement reporting. Although the campus comparison group was not comparable in relation to the campus poverty level, the current grouping process produced a fair and reliable system of comparison. The value added capacity of high performing campuses was equally rewarded in both the current and alternative processes.

However, an apprehension about using the order of dominance as used in the current system still exists. First, caution with the use of ethnicity as a sorting and comparison characteristic should be used. While policymakers purposefully design accountability systems to highlight disparity between ethnicities and emphasize the need for change, several political implications of this practice remain. While discussing the politics of reform, Cuban (1990, p. 16) once asked an appropriate question for this dilemma, "Are we dealing with the problem or the politics of the problem?" Though the state has a compelling reason to group by race in accountability systems, the difference in achievement levels among student groups makes the public uncomfortable. However, to ignore the achievement gap between different student groups would be an educational

injustice while establishing varying expectation for each ethnicity would be political suicide. Or in this case, do states have the wrong solutions to the right problems, or vice versa (Cuban, 1990)?

Second, some theorize (Jencks & Phillips, 1998) that grouping people by ethnicity produces psychologically negative reactions for minorities. In addition to dealing with negative stereotypes, people from minority groups must also deal with the fear of confirming such low expectations. This negative response has been termed stereotype vulnerability. Steele and Aronson (cited in Jencks & Phillips, 1998) proposed that as a result of being compared to whites, many minority students perform poorly on achievement tests. This stereotype threat is most prevalent for successful and academically strong African American students and may lead the students to withdraw. With each academic test, the black student fears confirming the society's negative stereotype of the past.

Another negative implication of using ethnicity as a sorting and comparison criteria is the potential carryover into curriculum and instruction. If stereotype vulnerability, as theorized by Steele and Aronson's (cited in Jencks & Phillips, 1998), also exists in the classroom, the impact becomes even greater as the child of minority is devalued by minimal expectations and relegated to lower level courses. If poverty, not ethnicity, is used, the negative stereotype and the detrimental legacy can be avoided.

The potential for harmful effects can be found in the Texas accountability system. Each campus's accreditation rating is based on the performance of all student and student groups of African American, Hispanic, white and economically disadvantaged. Due to

low TAAS performance, any one of the student groups could be responsible for a campus's low performance rating, thereby equating the campus low performance rating to one student group. In the Texas accountability system, the comparison of student performance by ethnicity could negatively affect minority student performance, and therefore, the campus performance. This comparison by ethnicity could prove to be counterproductive to the state's mission of excellence and equity for all students.

Although some empirical support for stereotype vulnerability exists, it is important to remember that it is only a theory. However, as reviewed in this study, sorting by poverty instead of ethnicity removes such negative stigmatization and any potential legal ramifications of a race-based system, because poverty cuts across all racial and ethnic groups. Additionally I found, sorting by poverty is equally reliable. In effect, substituting poverty for ethnicity as a primary sorter may have salutary consequences for both students and schools. The message that student performance is more indicative of educational opportunities than skin color is more palatable and equally reliable.

Suggestions for Further Research

Accountability is an intensely political issue facing educators. To achieve the goal of improved student performance, states have developed intricate information systems. With such information resources, much can be learned. TEA maintains a vast database of student performance. Although this study did not include a case study methodology, a study of Quartile 1 campuses compared to Quartile 4 campuses could reveal a profile of effectiveness useful in campus improvement. Additionally, no

longitudinal studies have been conducted using comparable improvement quartile placement as a research variable. A multi-year analysis of campuses consistently ranked in Quartile 1 or a case study of high performing or high gaining campuses could also provide additional information in campus improvement efforts.

An additional area for research is the impact of accountability reporting by ethnicity. In addition to dealing with negative stereotypes of lower mental ability, African Americans must also deal with the fear of confirming such low expectations (Jencks & Phillips, 1998). If the theory of stereotype vulnerability has value, the reporting of student performance by ethnicity may create unintentional detrimental pressure. The further review of stereotype vulnerability and its impact on various ethnicities could be valuable in the development of high quality accountability systems.

Because this study found the state comparable improvement a reliable comparison and the theory of stereotype vulnerability suggests that sorting by ethnicity has potential negative implications, a replication of this study with poverty as the primary but as the only sorting characteristic could review the impact of poverty more specifically. In this study, I used poverty as the primary sorting characteristic but completed the grouping process using the order of dominance of the remaining characteristics. If using only poverty produced similar results, the use of ethnicity could be eliminated altogether.

Although this study found the state comparable improvement a reliable comparison in terms of student performance, only elementary campuses were included in the analysis. Valuable information could be gained by the study of junior high or high school campuses. As this study focused on comparable improvement and poverty, many

unanswered questions remain on the grouping process regarding campus size and geographic location. As the race lines continue to blur, Clinchy's (2001, p. 494) "economic segregation" recognizes poverty status as the major characteristic that separates Americans rather than ethnicity. According to Clinchy, the high minority inner city and the predominantly white rural areas have higher incidents of extreme poverty and economically segregation than other locations. An attempt to construct new indicators or comparisons more comparable by size or geographic location could add to the findings of the current research.

In defining high performing students, TEA excludes TLI scores of the students that scored above 84 on TAAS. By applying the exclusions for high performing students, TEA is actually comparing the growth of the remaining students who did not score above a TLI of 84 to assign quartile placement. Although the comparable improvement process sorts campuses by total demographics, the comparison based on the demographics of the selected group of students whose TLI scores were actually included would also provide valuable information on campus comparisons.

In the review of research, teachers were more negative toward accountability measures than were administrators. Additionally, campus administrators were more negative than were district administrators. The closer the person was to the classroom, the more negative the person viewed accountability. A study of personal or professional motivational traits of educators could provide insight into why accountability systems are viewed as a strong instructional focus by some and a barrier to instruction by others. If

the role of accountability rewards is to motivate, how motivating do educators view this external force?

Conclusion

The challenge for any state accountability system is to provide valuable information about the instructional program in a fair and equitable manner. The fact that the data are publicly reported makes the data powerful and influential. The use of comparable improvement in the accountability and award system is a very important process that impacts all campuses in Texas. As a foundation of any accountability system, a state should reward the value added quality of effective instruction. The high performance of high poverty, high minority campuses emphasize the reality that when teachers have high expectations and effectively teach, children learn. The value added component, or the teacher and delivery portion, must be recognized and measured by a state accountability system.

In regard to comparisons, the idea of a system of differentiated expectations based on ethnicity would violate America's basic premise of equality and the effective school principle that all kids can learn. The use of poverty has and continues to be the only politically acceptable tool for differentiation. The federal government has and continues to recognize the impact of poverty, rather than ethnicity, as a measure to estimate instructional need and to generate supplementary funding. Somehow it is much more palatable to use poverty as the principal factor impacting student learning than ethnicity. The use of ethnicity to determine educational need would not be so acceptable. With the

diversification of America, it would be politically unacceptable to fund schools based upon the percentage of African American, Hispanic, or white students. And yet, the current system of comparable improvement does not recognize poverty as the most appropriate comparison for performance and growth. Using the most dominant characteristic to select the initial 100-campus pool, gives equal status to ethnicity, poverty, mobility, and language proficiency. Although the alternative process did not provide a substantial technical improvement, sorting by poverty may provide an immeasurable political improvement over the current process.

At the student level, it should matter not if the child is black or white. What must matter is what opportunities or experiences the child has missed due to lack of resources. Schools can and must focus on equality of educational opportunity. If the school is to level the playing field for all children, what task must be completed to level the field for the black or Hispanic child? And what if this black or Hispanic child comes from an affluent family? Because the most important factor that separates people is poverty, the school must respond to the child's lack of resources, not skin color. An affluent Hispanic campus has less in common with high poverty Hispanic campus than with another affluent campus, regardless of ethnicity. In this situation, a comparison based on poverty level would be the most valuable measure.

According to Cuban (1990), Americans view schools as the vehicle for social and individual change. Therefore, when a social problem arises, the school becomes the change agent. For poverty, the school has a clear mission to provide the necessary experiences and resources that the impoverished home was not able to provide. As

policymakers look to the school to solve the nation's social ills, education becomes the chosen path out of poverty for any poor child, black or white. The path remains the same.

REFERENCE LIST

Adams, N. (1995). Effects of AEIS on student academic achievement as perceived by Texas principals and teachers in the Region V ESC (Doctoral dissertation, Texas A & M University, 1995). Dissertation Abstracts International, 56, 3379.

Adams, E. R. (1994). The effects of cost, income, and socio-economic variables on Student scholastic aptitude scores (Doctoral dissertation, Ball State University, 1994). Dissertation Abstracts International, 55, 2276.

Alexander, N. A. (1998). The impact of curriculum standards on student performance: The case of New York State (Doctoral Dissertation, State University of New York at Albany, 1998). Dissertation Abstracts International, 59, 1770.

Allen, C.S. (1998). Metacognition, reading, and test taking of third graders. (Doctoral dissertation, Tennessee State University, 1998). Dissertation Abstracts International, 59, 3390.

Allen, H. H. (1991). A correlational study of achievement as measured by the total score on the Texas Academic Skills Program (TASP) reading test and achievement as measured by the total score on the Nelson-Denny reading test. (Doctoral Dissertation, University of Houston, 1991). Dissertation Abstracts International, 52, 2870.

Anrig, G. (1992, September). Testing and accountability. The American School Board Journal. 34-36.

Banks, K. (1994, April). Assessment's Conflicting Purposes, Conflicting Politics: Impact on Local School Systems. Paper presented at the annual Meeting of the American Educational Research Association, New Orleans, LA.

Ball, G. C. and Goldman, S. (1997). Improving education's productivity: Reexamining the system to get the schools we need. Phi Delta Kappan. November 1997. 228-232.

Balz, D. and Morin, R. (2000, January). 'Education Voters' Pose a Tough Test; Poll Shows Gore, Bush Face a Conflicted Group. The Washington Post. June 30, 2000.

Bechtel, D. (1997). Principal perceptions of accountability, autonomy, and assessment as part of school restructuring deal. (Doctoral dissertation, University of Pittsburgh, 1997). Dissertation Abstracts International, 58, 1992.

Beck, B. G. Q. (1995). State-mandated test score accountability: It's usefulness and meanings for school districts (Doctoral dissertation, University of Utah, 1995). Dissertation Abstracts International, 56, 4370.

Bennett, S. J. (1997). Empowering leadership: A multisite case study of school restructuring and accountability for student achievement (Doctoral dissertation, University of San Diego, 1997). Dissertation Abstracts International, 58, 654.

Bernauer, J. A. and Cress, K. (1997). How school communities can help redefine accountability assessment. Phi Delta Kappan. September. 1997. 71-75.

Berne, R. and Picus, L. O. (1994). Outcome Equity in Education. Corwin Press, Inc. Thousand Oakes, California. Editors.

Birt, P.J. (1998). Effects of performance-based accreditation probation status on student performance and teacher perceptions of an Indiana public school. (Doctoral dissertation, Indiana State University, 1998). Dissertation Abstracts International, 60, 1882.

Bonsangue, M. V. (1992). The effects of calculus workshop groups on minority achievement and persistence in mathematics, science, and engineering. (Doctoral dissertation, The Claremont Graduate School, 1992). Dissertation Abstracts International, 53, 3132.

Bourke, S. (1998, December). School Level Variables as Predictors of Individual Student Achievement. Paper presented at the annual meeting of the Australian Association of Research in Education, Adelaide, South Australia.

Bradley, R. (1998). School responses to the New York State report card (Doctoral dissertation, Syracuse University, 1998). Dissertation Abstracts International, 59, 2787.

Brown, M. E. (1998). A comparative study of the effectiveness of an extended year program for kindergartners. (Doctoral dissertation, George Washington University, 1998). Dissertation Abstracts International, 59, 490.

Brown, M. L. (1994). Effects of mobility on reading achievement as determined by performance-based assessment (Doctoral dissertation, University of La Verne, 1994). Dissertation Abstracts International, 56, 876.

Brown, P.R. (1990). Accountability in Public Education. Policy Briefs. Far West Laboratory. Number 14, 1-6.

Brownstein, R. (2000, September). National Perspective; New President Must Take Small Steps on Domestic Issues Before a Big Leap. Los Angeles Times. September 25, 2000.

Burke, W. M. (1999). A teacher-initiated and managed school improvement effort: Teachers' perceptions of its history and issues of implementation (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1999). Dissertation Abstracts International, 60, 3207.

Burstein, L., Oakes, J., & Guiton, G. (1992) Education indicators. Encyclopedia of Educational Research, (Vol. 2, pp. 410-418). New York: Macmillan.

Bush, B. L. (1999). Factors that support and impede full institutionalization of the middle school concept (Doctoral dissertation, University of Florida, 1999). Dissertation Abstracts International, 60, 4255.

Castellano, J. A. (1993). Perceptions of public school superintendents and principals Regarding the Illinois state board of educations' mandate on linking accountability to student performance and school improvement: The new Illinois Recognition system. (Doctoral dissertation, Northern Illinois University, 1993). Dissertation Abstracts International, 54, 1161.

Cauble, M. G. (1992). Selected purposes and quality indicators of Texas public school education and academic excellence indicators in Senate Bill 1 as perceived by selected Texas business and educational leaders. (Doctoral dissertation, Texas A&M University, 1992). Dissertation Abstracts International, 54, 36.

Cavazos, J. M. (1999). The instructional leadership of high school principals in successful Hispanic majority high schools. (Doctoral dissertation, University of Texas at Austin, 1999). Dissertation Abstracts International, 60, 4256.

Clifford, L. S. (1995). The impact of state mandated student performance assessments on instructional practices (Doctoral dissertation, University of Louiville, 1995). Dissertation Abstracts International 56, 4216.

Clinchy, E. (2001, March). Needed: A new educational civil rights movement. Phi Delta Kappan. 493-498.

Clinton, William J. (2000). Farewell Speech to the Nation. January 18. Speech.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfield, F.D., and Your, R.L. (1966). The equality of educational opportunities. Washington, DC: U.S. Department of Health Education and Welfare, U.S. Government Printing Office.

Collier, J. K. (1994). An educational production function analysis of selected Texas school districts (Achievement). (Doctoral dissertation, Texas Tech University, 1994). Dissertation Abstracts International, 55, 3360.

Connerly, W. (1999) Don't Discriminate. Wall Street Journal. Monday, October 4, 1999 Section A, p: 42.

Cormier, M. F. (1992). School personnel factors and community economic indicators as predictors of student performance on the Maine Educational Assessment Test. (Doctoral dissertation, Boston College, 1992). Dissertation Abstracts International, 54, 154.

Council of Economic Advisers for the President's Initiative on Race. (1998). Changing America. Indicators of Social and Economic Well-Being by Race and Hispanic Origin. (Government Printing Office ISBN 0-16-049616-0). Washington, D.C. Author.

Cuban, L. (1990, January). Reforming again, again, and again. Educational Researcher.19. 1. 10-21.

Cuban, L. (1992). The corporate myth of reforming public schools. Phi Delta Kappan. 74 (2), 157-159.

Darling-Hammond, L. (1989). Accountability for professional practice. Teachers College Record. Volume 91. Number 1. 58-80.

Darling-Hammond, L. (1992). Educational indicators and enlightened policy. Educational Policy. 235-265.

Darling-Hammond, L. and McLaughlin, M.W. (1995, April). Policies that support professional development in an era of reform. Phi Delta Kappan. 597-604.

Davies, A. and Williams, P. (1997, September). Accountability: Issues, possibilities, and guiding questions for districtwide assessment of student learning. Phi Delta Kappan. 7679.

DeBray, E. (1999, April). Incentives in States' Educational Accountability Systems: Is the Assumption of Continuous Improvement Included? Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Decker, L. and Decker, V. (1988). Home/School/Community Involvement. American Association of School Administrators, Arlington, Virginia. 1988.

DeMouldin, D. F. and Kendall, R. (1993). The administrative role in an accountability network: A developmental conceptualization. Journal of School Leadership. 3, November. 688-698.

Dollinger, C. Z. (1997). Excessive absenteeism: Comprehending the issues involved in high school students' decisions regarding non-attendance. (Doctoral dissertation, University of Sarasota, 1997). Dissertation Abstracts International, 58, 4607.

Edington, E. D. and Martellaro, H. C. (1984, April). Variables Affecting Academic Achievement in New Mexico Schools. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, Louisiana.

Edmonds, Ronald. (1979). Effective schools for the urban poor. Educational Leadership. October, 15-24.

Eineman, T. A. (1992). The effects of retention on elementary school student performance as measured by legally mandated standardized achievement tests. (Doctoral dissertation, Ball State University, 1992). Dissertation Abstracts International, 54, 806.

England, D. C. (1993). The impact on the success of high-risk students of placement policies established by Texas higher education institutions in the implementation of the Texas Academic Skills Program. (Doctoral dissertation, Texas A&M University, 1993). Dissertation Abstracts International, 54, 2038.

Fetler, M. (1990). California's Role in School Accountability. Policy Briefs. Far West Laboratory. Special Insert. 1-2.

Feuer, M.J. and Kober, N. (1995, March). Anticipating Goals 2000. Standards, Assessment, and Public Policy. (Summary of a Workshop). National Research Council, Commission on Behavioral and Social Science and Education. Washington, D.C.

Fortune, R. F. (1979). The effects of race and socio-economic status on student achievement. (Doctoral dissertation, University of California, 1979). Dissertation Abstracts International, 41, 147.

Franklin, B. J. and Crone, L. J. (1992, November). School Accountability: Predictors and Indicators of Louisiana School Effectiveness. Paper presented at the annual meeting of the Mid-South Educational Association, Knoxville, Tennessee.

Gaines, G. R. and Cornett, L. M. (1992). School Accountability Reports: Lessons Learned in SREB States. (Policy Paper). Atlanta, Georgia. Southern Regional Education Board.

Garn, G. A. (1998). The accountability system for Arizona charter schools (Doctoral dissertation, Arizona State University, 1998). Dissertation Abstracts International, 59, 1410.

Haetinger, J. Interview. Texas Education Agency. Research and Development. July 2000.

Harrington-Lueker, D. (1998, June). Raising the Bar. Now local school districts are accountable for results. The American School Board Journal. 17-21.

Hart, R. D. (1999). The impact of high stakes accountability on high school curriculum (Doctoral dissertation, Texas A & M University- Commerce, 1999). Dissertation Abstracts International, 60, 3854.

Heffron, J. L. (1998). Strategies used by principals to establish conditions that foster improved student performance in site-based managed schools and the strategies used by district office administrators to support the principals. (Doctoral dissertation, University of La Verne, 1998). Dissertation Abstracts International, 59, 2796.

Helfand, D. (2000, September). Campaign 2000. Test Scores Contradict; 'Recession' Claim; Schools; Students have shown slow improvement in two decades despite Bush's assertions, experts say. Los Angeles Times. September 27, 2000.

Hemberger-Coffey, J. (1991). Student achievement, community socioeconomic characteristics and sategories of educational expenditure in Connecticut public schools (Doctoral dissertation, University of Bridgeport, 1991). Dissertation Abstracts International, 52, 3140.

Henry, T. (2000, September). Learning to tell them apart Gore, Bush have same goals: Better education. USA TODAY. September 26, 2000.

Henry, G. T. (1996, September). Community accountability: A theory of information, accountability, and school improvement. Phi Delta Kappan. 85-90.

Herrington, C. (1993, April). Accountability, Invisibility, and the Politics of Numbers: School Report Cards and Race. Paper presented at the Annual Meeting of American Educational Research Association. Atlanta, Georgia.

Henson, R., and Smith, A.D. (2000, Summer). State of the art in statistical significance and effect size reporting: A review of the APA Task Force Report and current trends. Journal of Research and Development in Education. 33, 285-96.

Hill, A. N. (1999). Examining the effects of school climate, student background, teacher expectations, and school-community integration on academic achievement among high poverty urban schools (Doctoral dissertation, University of Houston, 1999). Dissertation Abstracts International, 60, 1018.

Hinkle, D.E., Wiersma, W., and Jurs, S. G. (1998). Applied Statistics for the Behavioral Sciences (4th ed.). Boston, Massachusetts: Houghton Mifflin Company.

Holman, L. J. (1993). Predictors of student performance on a high-stakes criterion-referenced test in a United States-Mexico border Area. (Doctoral dissertation, New Mexico State University, 1993). Dissertation Abstracts International, 54, 2824.

Howard, T. L (1998). A qualitative study: Teachers' perceptions of North Carolina's Accountability, Basics and Control plan. (Doctoral dissertation, North Carolina State University, 1998). Dissertation Abstracts International, 59, 3703.

Hunt, A. R. (1999, December 23). Politics and People: Yuletide is for the Poor, Too. The Wall Street Journal, Sec. A., 19.

Jaeger, R. M., Gorney, B. E., & Johnson, R. (1994, October). The other kind of report card: When schools are graded. Educational Leadership. 42-45.

Jencks, C. and Phillips, M. (1998). Editors. The Black-White Test Score Gap. Brookings Institution Press. Washington D.C.

Johnson, P. A. (1997). Reframing restructuring: Power relationship in arts-based schools. (Doctoral dissertation, University of Denver, 1997). Dissertation Abstracts International, 58, 1163.

Kane, M. (1993). The Validity of Perforamnce Standards. National Assessment Governing Board, Evaluation Report. June 19, 1993. .Washington, DC.

Kaminski, L. R. G. (1999). Here today and gone tomorrow: The impact of student mobility on learning. (Doctoral dissertation, Harvard University, 1997). Dissertation Abstracts International, 60, 1997.

Kelley, C. (1998). The Kentucky School-Based Performance Award Program: School-Level Effects. Educational Policy, 12, 305-324.

Labaree, F. (2000, September). Resisting educational standards. Phi Delta Kappan. p.28-33.

Lamberson, J. E. (1989). Educational accountability as assessed by selected 1986-1987 Illinois school report card achievement, school and student variables (Doctoral dissertation, Illinois State University, 1989). Dissertation Abstracts International, 50, 1159.

Landry, D. A. (1997). An investigation of selected variables to predict student performance on the National Council Licensure Examination for Registered Nurses (NCLERN) in one baccalaureate degree nursing program. (Doctoral dissertation, West Virginia University, 1997). Dissertation Abstracts International, 58, 4511.

LeTendre, M. J. (2000). U.S. Department of Education Director of Compensatory Education. Keynote Speech. National Title 1 Conference. Phoenix. Arizona. January 22. Speech.

Lezotte, L.W., (1986, April). School Effectiveness: Reflections and Future Direction. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA.

Lieberman, A. (1988). Editor. Building a Professional Culture in Schools. Teachers College Press, Teachers College. Columbia University. New York, New York.

MacDonald, T. P. (1996). Determining the relationship of selected district and student demographics with eighth-grade achievement in Alaska (Doctoral dissertation, University of Southern Mississippi, 1996). Dissertation Abstracts International, 57, 4209.

Mayer, S. E. (1997). What Money Can't Buy. Family Income and Children's Life Chances. Harvard University Press. Cambridge, Massachusetts.

Maume, M. O. (1998). Secondary control: Examining the influence of school restructuring on high school delinquency. (Doctoral dissertation, Louisiana State University and Agricultural and Mechanical College, 1998). Dissertation Abstracts International, 59, 2195.

McClure, R. D. (1999). Low performing to exemplary: Successful change in one elementary school. (Doctoral dissertation, University of Texas at Austin, 1999). Dissertation Abstracts International, 60, 3221.

McGinn, D. (1999, September 6). The Big Score. Newsweek. 46-51.

McLaughlin, M.W (1991, November). Test-Based accountability as a reform strategy. Phi Delta Kappan, 248-251.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. Education Researcher, 18 (2), 5-11.

National Center for Excellence in Education. (1983). A Nation at Risk: The imperative for Educational Reform. Washington, D.C.: Author.

Neal, T. M. (2000). Hitting Harder on Education; Bush Alleges '7 Tears of Stagnancy' and Pledges Change. The Washington Post. August 31, 2000.

Newmann, F. M., King, M.B., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. Harvard Educational Review. 67, 41-73.

Nine, G. S. (1990). Measuring teacher effectiveness: Process (Performance Level Placement) versus product (Standardized Student achievement). (Doctoral dissertation, Northern Arizona University, 1990). Dissertation Abstracts International, 51, 1588.

Odden, A. (1992). School finance the 1990's. Phi Delta Kappan, 73 (6), 455-461.

Oldham, B.R., (1993, April). High Stakes Assessment: A Local District Perspective. Paper presented at the Annual Meeting of the American Educational Research Association. Atlanta, Georgia.

Perkins, K. E. (1992). The relationship between unit school district instructional expenditure per pupil and academic achievement (Doctoral dissertation, Southern Illinois University, 1992). Dissertation Abstracts International, 53, 2682.

Pierre, P. N. (1994) Exploration of how home and parental characteristics support achievement and effective school functioning of adolescents (Doctoral dissertation, Emory University, 1994). Dissertation Abstracts International, 55, 1193.

Plecki, M. L. (1991). The relationship of elementary school size and student achievement inCalifornia public schools. (Doctoral dissertation, University of California, Berkeley, 1991). Dissertation Abstracts International, 53, 1351.

Powell, C. A. (1997). The nature of effective responses to state algebra I policy at high poverty Texas high schools (Doctoral dissertation, University of Texas at Austin, 1997). Dissertation Abstracts International, 58, 2576.

Ramirez, A., McClanahan, R. (1992, April). Reporting to the public. The American School Board Journal. 33-35.

Rathbone, S. (1988). A process model for the determination of indicators of school success (Doctoral dissertation, University of Delaware, 1988). Dissertation Abstracts International, 50, 45.

Renaldi, R. P. (1992). The New Jersey school report card: A case study of the initiation of a state-level accountability system for local public schools (Doctoral dissertation, Pennsylvania State University, 1992). Dissertation Abstracts International, 53, 2635.

Richards, C. (1988, March). Indicators and three types of educational monitoring systems: Implications for design. Phi Delta Kappan. 495-499.

Ricotta, F. J. (1987) A Study of Colorado Education Accountabilty Act of 1971 mandates as implemented by the state policy makers (Doctoral dissertation, Univeristy of Colorado at Boulder, 1987). Dissertation Abstracts International, 48, 1080.

Riegel, L. (1992, April). Texas Campus Report Cards: The Story of One Blind Man and Six Elephants. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Saban, A. (1997, July). Emerging Themes of National Educational Reform. International Journal of Educational Reform. v. 6, n. 3, 349-56.

San Diego City Schools. School Services Division, Area VI (1993, March 23). Report andRecommendations of the Student Achievement Accountability Committee (SAAC). San Diego, California.

San Miguel, T. (1996). The Influence of the state-mandated accountability system on the school improvement process in selected Texas elementary schools (Doctoral dissertation, University of Texas at Austin, 1996 Dissertation Abstracts International, 58, 52.

School Improvement. (2001). Favorite Quotes. Website. http://mayacad.org/Scho_improve_savings.html.

Seidenfeld, A. R. (1991). The effects of four spelling study strategies and student performance. (Doctoral dissertation, University of Denver, 1991). Dissertation Abstracts International, 52, 2446.

Smith, M. O'Connell (1999). How the apportionment of district money is associated with differences in the grade 8 mathematics and reading achievement scores for Minnesota students (Doctoral dissertation, University of Minnesota, 1999). Dissertation Abstracts International, 60, 1908.

Soltz, D. F. (1991). An analysis of variables contributing to the effectiveness of test-Wiseness preparation. (Doctoral dissertation, University of Colorado at Boulder, 1991). Dissertation Abstract International, 52, 2477.

Southern Regional Education Board. (SREB) (1995). Linking Education Report Cards and Local School Improvement. Educational Benchmark. Goals for Education. Altanta, Georgia: Author.

Stout, R. T. (1990). Accountability in Arizona. Policy Briefs. Far West Laboratory. Special Insert. 1-2.

Streshly, W.A., & Newcomer, L. (1994, March). Managing Change with Accountability: A Challenge for Educators. NASSP Bulletin, 62-68.

Student Achievement Accountability Committee Report and Recommendations (SAAC). (1993). Executive Summary. San Diego City Schools, California: Author.

Texas Education Agency. (2000). 2000 Accountability Manual: The 2000 Accountability Rating System for Texas Public Schools and School Districts. Office of Policy Planning and Research. Austin, Texas.

Texas Education Agency. (2000) Glossary. Office of Policy Planning and Research, Austin,Texas.

Texas Education Agency. (2000) Frequently Asked Questions: Texas Successful Schools Award System. Website. http://www.tea.state.us/TSSAS/TSSASfaq.htlm.

Texas Education Agency. (2000). Pocket Edition. (TEA). Austin, Texas.

Texas Education Agency. (1999). Texas Education Code. Austin, Texas. West Group.

Texas Education Agency. (2000, August). The Texas Successful Schools Study: Quality Education for Limited English Proficient Students. Research Study. Austin, Texas.

Theobald, P., and Mills, E. (1995, February). Accountability and the struggle over what counts. Phi Delta Kappan. 462-466.

Thompson,F. T. III. (1996). Relationships between selected environmental characteristics, neighborhood type, and school achievement. (Doctoral dissertation, University of Nebraska at Omaha, 1996). Dissertation Abstract International, 57, 4696.

Toner, R. (2000, August). Platform Strives to Reach Right and Center. The New York Times. August 1, 2000.

United States Department of Education. A Nation Responds (1984). Washington, DC: Government Printing Office.

United States Department of Education. (1999). Dispelling the Myth: High Poverty Schools Exceeding Expectations. Report of the Education Trust in cooperation with the Council of Chief State School Officers. U. S. Department of Education, Washington. D.C. 1999.

United States Department of Education. (2000). Final Allocations. Webpage. http:www.ed.gov:80/offices/OESCEL/compreform/scrdfly2000final.html. Final Allocations. United States Department of Education. Washington DC.

Webb, M. W. (1995, April). Policy Considerations in Developing Standards and Assessments for Large, Diverse School Districts. Paper presented at the annual meeting of the National Council on Measurement in Education. San Francisco, CA.

Webster, B. (2001, February). Statewide Numbers of Economically disadvantaged Students by Ethnicity. Report. Austin, Texas.

Welch, N. (1995). The Kansas experience: A state-wide assessment and recommendation for community-based nutrition internships. (Doctoral dissertation, Kansas State University, 1995). Dissertation Abstracts International, 56, 3134.

Whitus, E. F. (1996). The predictive effects of the mathematics portion of the Texas Assessment of Academic Skills Test and other variables on the mathematics portion of the Texas Academic Skills Program. (Doctoral dissertation, Texas A&M University-Commerce, 1995). Dissertation Abstracts International, 58, 713.

White, E. M. (1994). Teaching and Assessing Writing. Jossey-Bass Publisher. San Francisco. Second Edition.

Wiggins, G.P. (1993). Assessing Student Performance. Exploring the Purpose and Limits of Testing. Jossey-Bass Publishers. San Francisco. First Edition.

Wiggins, G.P. (1998). Educative Assessment. Designing Assessments to Inform and Improve Student Performance. Jossey-Bass Publishers. San Francisco. First Edition.

Wilburn, S. D. (1996). A qualitative study of legislated assessment policy in Tennessee (Doctoral dissertation, University of Tennessee, 1997). Dissertation Abstracts International, 58, 2020.

Yong, R. S. (1987). The impact of wealth and size on selected accountability indicators of Illinois school districts (Doctoral dissertation, Illinois State University, 1987). Dissertation Abstracts International, 48, 537.

Young, M. D. (1996, April). Understanding family involvement in secondary schools: Lessons learned from Texas borderland schools. Paper presented at the annual meeting of the American Educaitonal Reseach Association (pp. 1-30). New York, NY.

Zankofski, D. A. (1999). The effects of demographic and background variables on student performance in mathematics for general education. (Doctoral dissertation, George Mason University, 1999). Dissertation Abstracts International, 59, 4040.